

Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego na studiach na kierunkach Elektronika i Telekomunikacja oraz Teleinformatyka.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

Wprowadzenie do multimediów

Zajęcia prowadzone przez:

Instytut Telekomunikacji Multimedialnej



- ***Wykład:***

prof. dr hab. inż. Marek Domański

- pokój 104, ul. Polanka 3, tel. 66 53 901

- e-mail:

marek.domanski@put.poznan.pl

- ***Konsultacje:***

- wtorki, godz. 15.00 – 15.45

- piątki, godz. 15.00 – 15.45

Wykłady

- **Prowadzenie notatek**
- **Zadawanie pytań**
- **Zasady obowiązujące podczas wykładów**

Telefony komórkowe i komputery są wyłączone i schowane.

- **Korzystanie z tabletu lub laptopa do notowania wymaga indywidualnej zgody wykładowcy uzyskanej odpowiednio wcześniej.**
- **Sprawdzanie obecności**

Nakład pracy na ten przedmiot

- **Ogólne zasady: 1 punkt ECTS to średnio ok. 30 godzin pracy studenta.
Ten przedmiot odpowiada 5 punktom ECTS.**
- **Zakłada się, że praca samodzielna pochłania dodatkowo mniej więcej tyle samo czasu, co udział w zajęciach.**
- **Te szacunki odpowiadają systematycznej pracy studenta. Jeżeli jest on przygotowany do zajęć, aktywnie w nich uczestniczy, to wynosi istotną wiedzę z wykładów i ćwiczeń, więc wystarcza poświęcenie ok. 75 godzin (średnio ok. 5 godzin tygodniowo) samodzielnej pracy.**

Nakład pracy na ten przedmiot

- **Wymienione dotychczas zasady dotyczą studentów, którzy mają dobrze opanowane przedmioty stanowiące wymagania wstępne, czyli oprócz analizy matematycznej, to są: teoria sygnałów, teoria systemów, cyfrowe przetwarzanie sygnałów, wprowadzenie do telekomunikacji, podstawy informatyki.**
- **Pozostali muszą poświęcić na ten przedmiot ZNACZNIE WIĘCEJ CZASU PRACY SAMODZIELNEJ (niż to wynika z liczby punktów ECTS), by uzupełnić zaległości.**

Jak uczymy się tego przedmiotu

- **NIE PAMIĘCIOWO !!!**
- Stosunkowo niewiele trzeba nauczyć się na pamięć.
- Staramy się zrozumieć zasady i metody, a następnie staramy się nauczyć ich stosowania.
- Pracujemy **SYSTEMATYCZNIE** przez cały semestr.
- Przedmiot jest bardzo ciekawy i jest głęboko zakorzeniony w praktycznej działalności.

Typowe błędy w nauce tego przedmiotu popelniane przez studentów

- **Niesystematyczna praca:**
Nieregularne uczęszczanie na wykłady powoduje, że uczestnictwo w kolejnych wykładach niewiele daje.
Brak przygotowania do ćwiczeń laboratoryjnych powoduje, że uczestnictwo w nich przynosi niewiele pożytku studentowi.
- **Rozpaczliwa nauka przed egzaminem:**
Próbuje się w ciągu paru dni nadrobić zaległości całego semestru.

Typowe błędy w nauce tego przedmiotu popelniane przez studentów

- **Typowa gorączkowa nauka przed egzaminem:**
 - Wzrokowa nauka polegająca na mechanicznym zapamiętywaniu treści przezroczy ilustrujących wykłady, bez wnikania (z braku czasu) w merytoryczne treści.
 - Uczenie się na pamięć rozwiązań zadań egzaminacyjnych z poprzednich lat.
 - Wkuwanie wyglądu wykresów lub schematów bez zrozumienia ich znaczenia.

Egzamin

- Egzamin ma formę pisemną i/lub ustną.
- Formę (ustny/pisemny/obie) egzaminu ustala egzaminator w zależności od możliwości organizacyjnych i innych okoliczności przeprowadzenia egzaminu.

Organizacja egzaminu pisemnego i ustnego:

- **Należy przynieść ze sobą:**
 - kilka kartek papieru,
 - przybory do pisania i sporządzania wykresów (w tym zapasowe),
 - kalkulator (z funkcjami trygonometrycznymi i wykładniczymi) .
- **Korzystanie w trakcie egzaminu z materiałów pomocniczych, tablic, notatek, skryptów, książek oraz programowalnych kalkulatorów i wszelkich komputerów (palmtopów, laptopów, tabletów) jest surowo zakazane. *Korzystanie z takich materiałów jest traktowane jako próba oszustwa.***
- **Telefony komórkowe muszą być wyłączane przed wejściem do sali egzaminacyjnej. Telefony komórkowe muszą być wyłączone i schowane w torbach aż do opuszczenia sali egzaminacyjnej. *Korzystanie ze smartfonów jest traktowane jako próba oszustwa. Dotyczy to także posiadania słuchawek w uszach lub na stole.***
- **Spóźnienie może uniemożliwić zdawanie egzaminu w danym terminie.**

Ocena prac egzaminacyjnych

- Podstawę oceny stanowią rozwiązania problemów egzaminacyjnych.
- Dodatkowe punkty można uzyskać za systematyczną obecność na wykładach.

Parę porad – nie wierzcie mitom

Parę mitów:

Bardzo trudny egzamin – nieprawda: Często liczba ocen dobrych i bardzo dobrych przewyższa liczbę ocen dostatecznych. Kto w ogóle zabrał się do nauki w sensowny sposób, ten z łatwością jest w stanie uzyskać dobrą ocenę.

Mnóstwo nauki pamięciowej – nieprawda: Przedmiot dotyczy praktycznej działalności technicznej – ta jest zawsze logiczna i uzasadniona. Oczywiście jest trochę faktów, które powinno się znać, ale one zazwyczaj wynikają z uwarunkowań technicznych, fizycznych, ekonomicznych lub społecznych, co istotnie ułatwia ich zapamiętanie.

Parę porad – nie wierzcie mitom

Kolejny mit: Mnóstwo schematów blokowych

do wykucia – nieprawda: Właściwie pojawiają się schematy trzech koderów i odpowiadających im dekodeków. Schematy opisują algorytmy, które należy zrozumieć. Na wykładach te same schematy pojawiają się wielokrotnie: czasem pozbawione wybranych bloków, czasem z podkreśleniem aktualnie omawianej operacji przedstawionej pewnym blokiem itp.

Jeżeli ktoś zna algorytm, to narysuje schemat blokowy. Każdy logiczny sposób jego narysowania jest akceptowany na egzaminie. Dla studenta rozumiejącego istotę działania algorytmu nie stanowi to trudności.

Dla studenta, który do egzaminu przygotowuje się starając się zapamiętać treść kilkuset przezroczy (23 wykłady), nauka staje się zmarnowanym czasem i traumatycznym przeżyciem nie prowadzącym do uzyskania użytecznej wiedzy i umiejętności.

Materiały do wykładów

<http://www.multimedia.edu.pl>

Aktualne: ekursy.put.poznan.pl

Zagadnienia

1. **Telekomunikacja multimedialna**
2. **Obraz cyfrowy**
 - reprezentacja i percepcja
3. **Podstawy cyfrowego przetwarzania obrazów**
4. **Kompresja obrazów – techniki**
5. **Kompresja sekwencji wizyjnych**
 - techniki
6. **Dźwięk cyfrowy**
 - właściwości, percepcja, reprezentacja i kompresja
7. **Sygnał mowy – właściwości, percepcja i kompresja**
8. **Przesyłanie treści multimedialnych**

Literatura w języku polskim

Obraz:

- **M. Domański, Obraz cyfrowy, WKiŁ, Warszawa 2010.**
- **D. Karwowski – Zrozumieć Kompresję Obrazu, 2019,**
https://www.zrozumieckompresje.pl/wp-content/uploads/DKarwowski_ZrozumieckompresjeObrazu.pdf .

Dźwięk:

- **A. Czyżewski, Dźwięk cyfrowy, AOW Exit, Warszawa 2001.**
- **W. Butryn, Dźwięk cyfrowy, WKiŁ, Warszawa 2001.**

1. Wprowadzenie

1.1. Telekomunikacja multimedialna

Multimedia

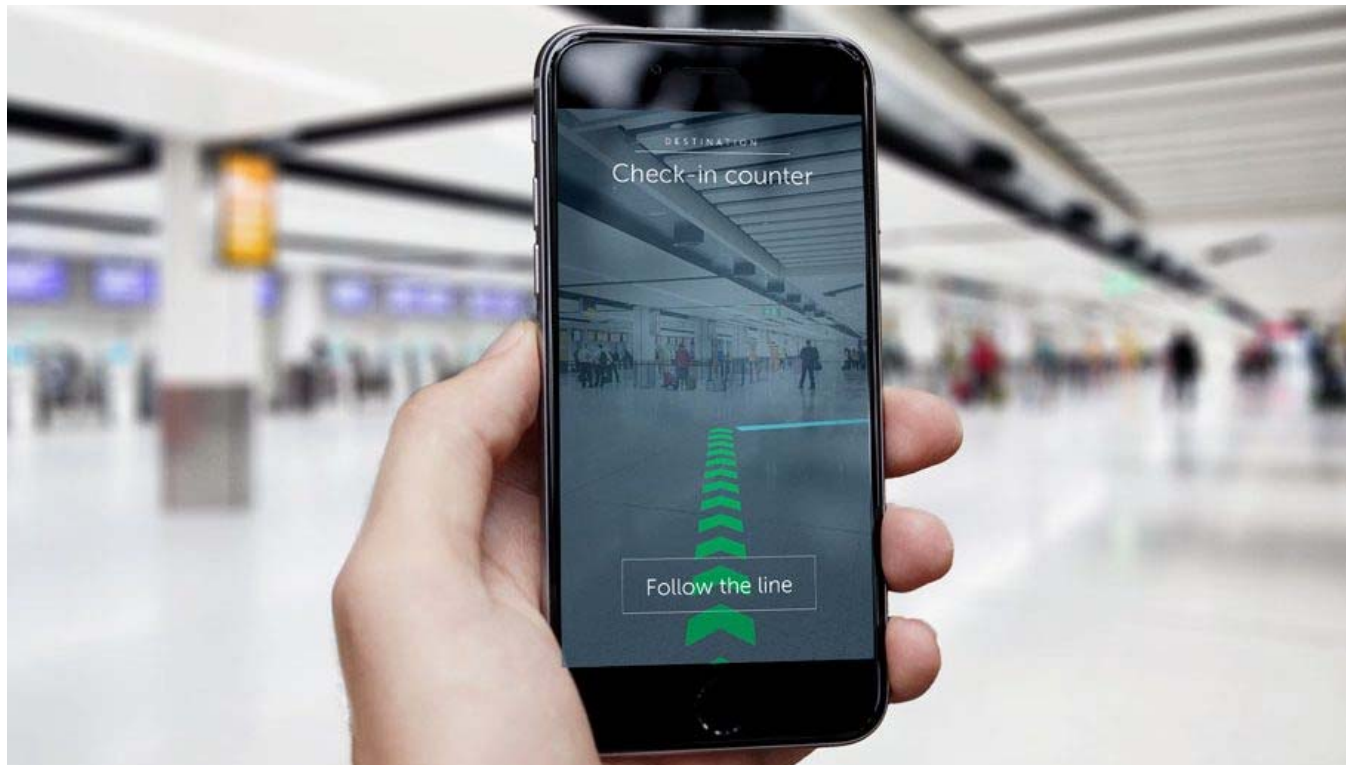
Pojęcie odnoszące się do komunikacji za pomocą kilku typów danych:

- **Tekst** (*text*)
- **Obraz (obraz nieruchomy)** (*image*)
- **Wizja (sekwencje wizyjne)** (*video*)
- **Grafika** (*graphics*)
- **Fonia (dźwięk)** (*audio*)
- **Mowa** (*speech*)
- **Nowe media** (*new media*)

Wzbogacona rzeczywistość

Augmented reality

Treść wygenerowana za pomocą komputera
nałożona na treść rzeczywistą



Główny obszar zastosowań: profesjonalne

Wirtualna rzeczywistość

Virtual reality

Treści sprawiające wrażenie fizycznej obecności w wirtualnym otoczeniu

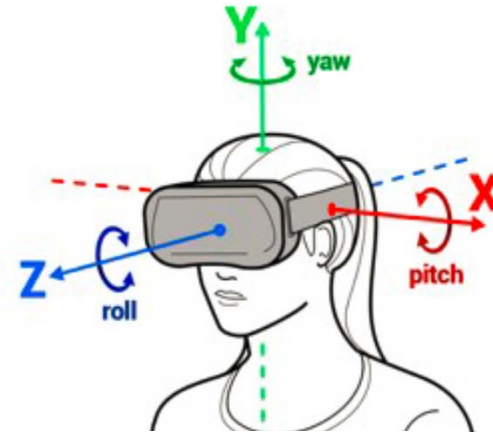


Główny obszar zastosowań: rozrywka

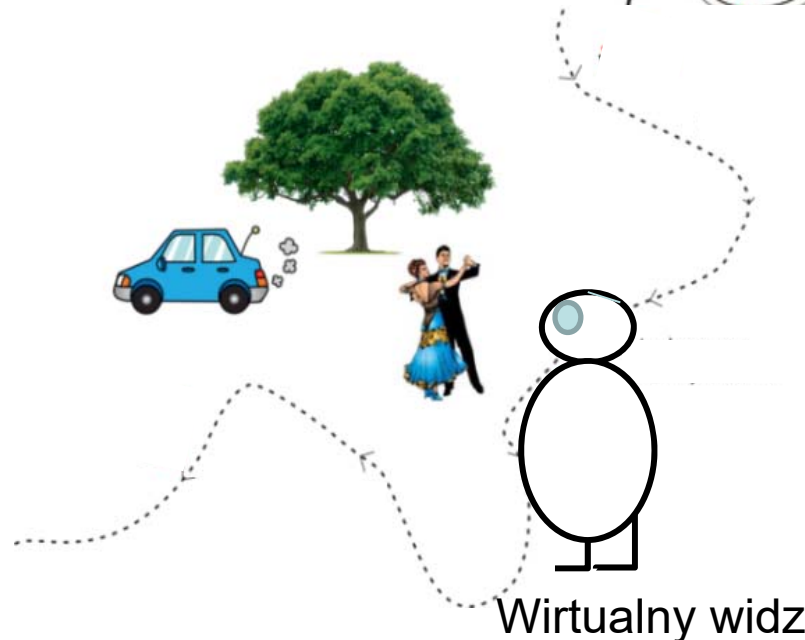
Media wszechogarniające

Immersive media

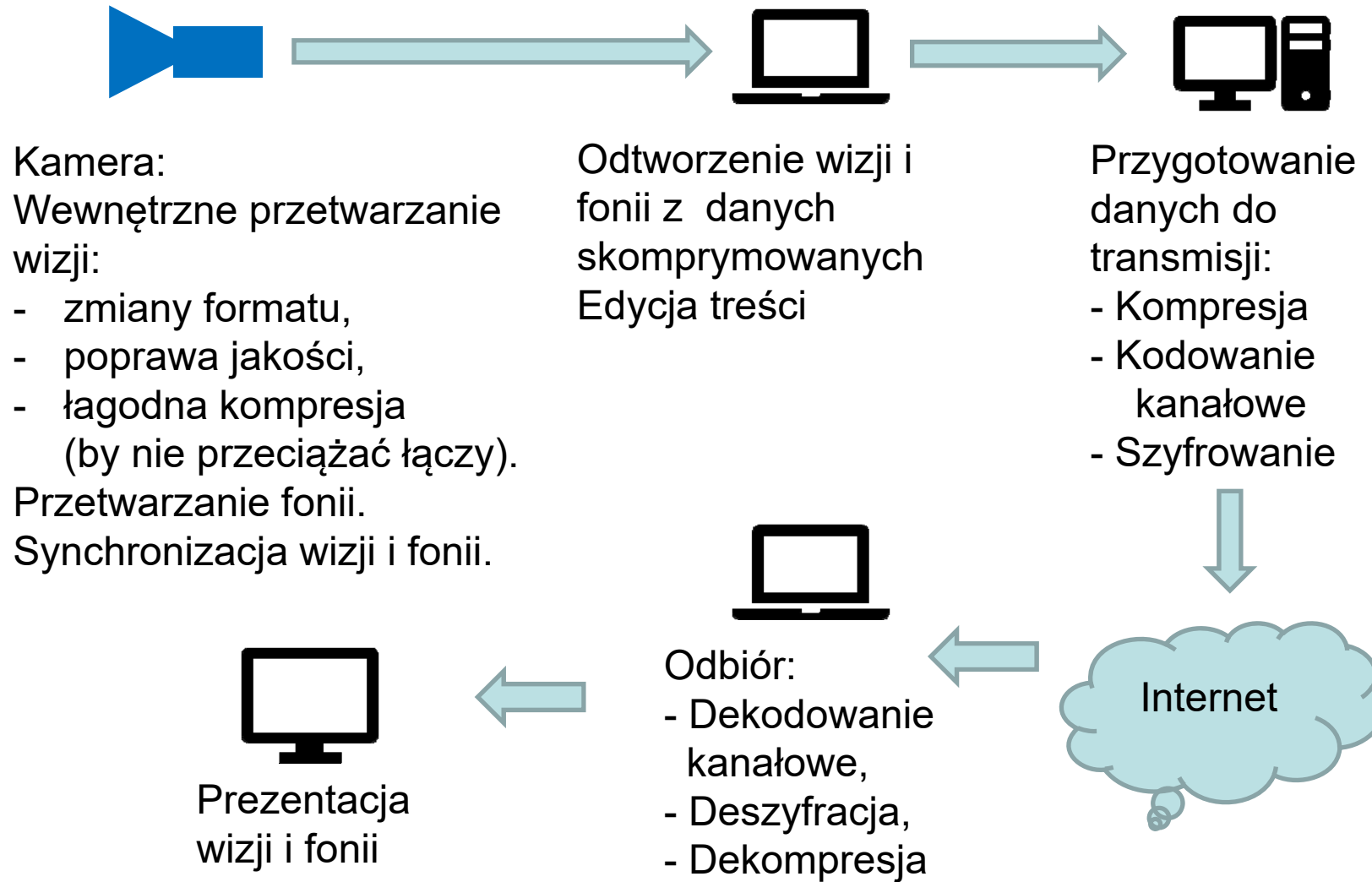
- **Obraz dookólny,
wszechkierunkowy**
- *Omnidirectional video
(image), 360° video*



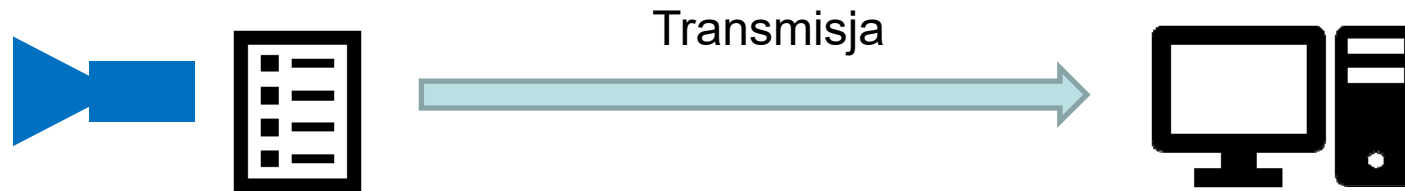
- **Wirtualna
nawigacja**
- *Virtual
navigation*



Przesyłanie filmów przez internet



Systemy dozoru wizyjnego



Inteligentna kamera

z przetwarzaniem wizji:

wykrywanie obiektów i zdarzeń.

Przygotowanie danych do transmisji:

- Kompresja
- Kodowanie kanałowe
- Szyfrowanie

Centrum dozoru:

Odbiór:

- Dekodowanie kanałowe,
- Deszyfracja,
- Dekompresja

Automatyczna analiza

+ prezentacja dla obsługi

Archiwizacja

1.2. Dane multimedialne

Mediadane

- reprezentacje zdjęć, filmów, utworów muzycznych itd.

Nieskompresowane (nieskomprymowane)

- wartości próbek.

Skompresowane (skomprymowane)

- reprezentacje wyliczone z próbek, próbki mogą być wyznaczane z tych reprezentacji.

Metadane

- dodatkowe dane opisujące medium (mediadane).

Tworzone przez ludzi, generowane automatycznie,
generowane półautomatycznie.

Wysokiego poziomu: np.: tytuł, autor, rok publikacji itd.,.

Niskiego poziomu: np.: barwa dominująca, uproszczony kształt,
charakter ruchu, rytm w utworze muzycznym itd.

Metadane są przydatne do poszukiwania obiektów multimedialnych.

- **Przesyłałyśmy coraz więcej obrazów**
- **Coraz lepsza jakość –
Nowe formaty obrazu:
np. UHDTV: 3840 x 2160 lub 7680 x 4320
Wizja 360°**
- **Udział wizji w całym ruchu rośnie**

Wizja

- **Globalnie: wizja IP tworzy ok. 82 % całego ruchu IP (konsumenckiego oraz profesjonalnego) w roku 2022; wzrost z 73 % w roku 2016.**
- **Średni roczny przyrost ruchu wizyjnego IP: 26 %.**
- **Całkowity ruch wizyjny: w latach 2016 - 2021 wzrost 4-krotny, średni roczny przyrost 31 %.**

**Telekomunikacja staje się
dziedziną techniki zajmującą się
głównie przesyłaniem obrazu**



Znaczenie kompresji

1.3. Kompresja danych multimedialnych

Motywacja:

Zazwyczaj duża objętość danych multimedialnych

Np. sygnał obrazu standardowej telewizji cyfrowej:

ok. 125 Mb/s

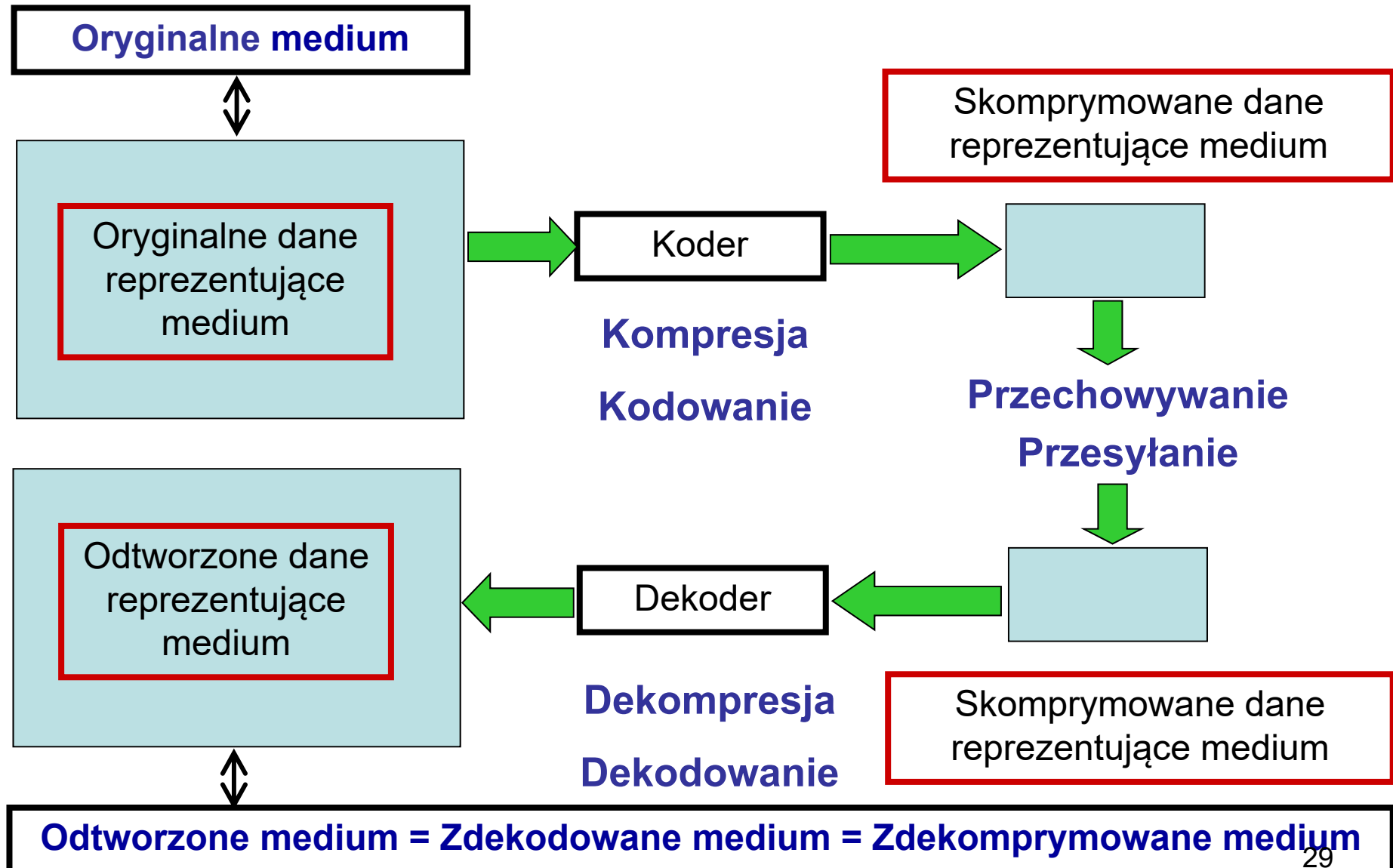
Po kompresji nawet <0.5 Mb/s

Sygnał obrazu telewizji cyfrowej wysokiej jakości (HDTV)

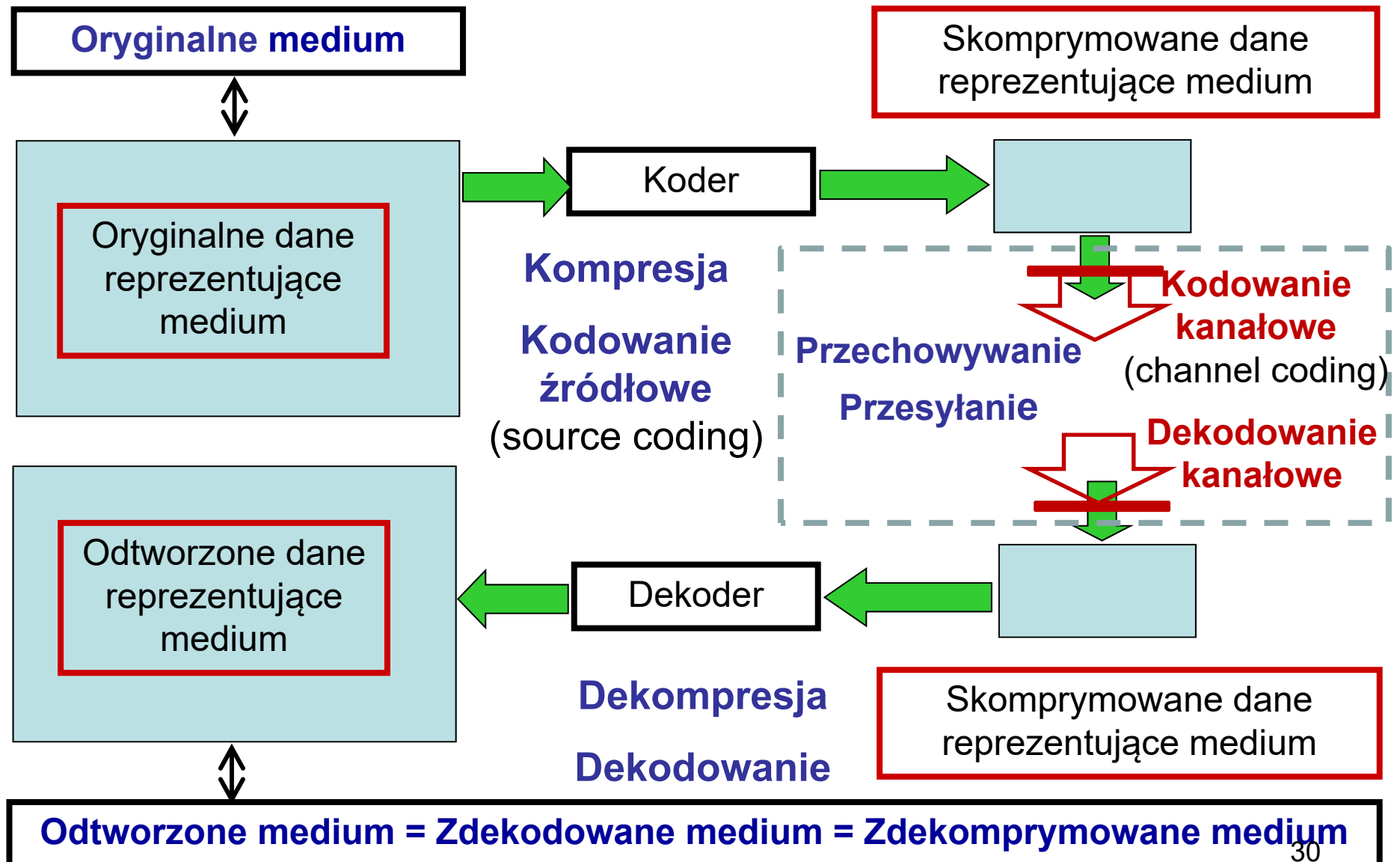
rzędu 620 Mb/s

Po kompresji nawet <2 Mb/s

Kompresja (*compression*)



Kompresja (*compression*)



Znaczenie kompresji

- **SDTV nieskompresowana wizja**
124 Mb/s
- **Prosta kompresja**
DV
25 Mb/s
- **Klasyczna kompresja (stara)**
MPEG-2
~ 3,5 Mb/s
- **Standardowa kompresja**
AVC (H.264)
~ 1,5 Mb/s
- **Nowsza technika kompresji**
HEVC
~ 0,8 Mb/s
- **Najnowsza technika kompresji**
VVC
~ 0,5 Mb/s

1.4. Normalizacja w zakresie multimedialów (*Standardization in multimedia*)

Cel:

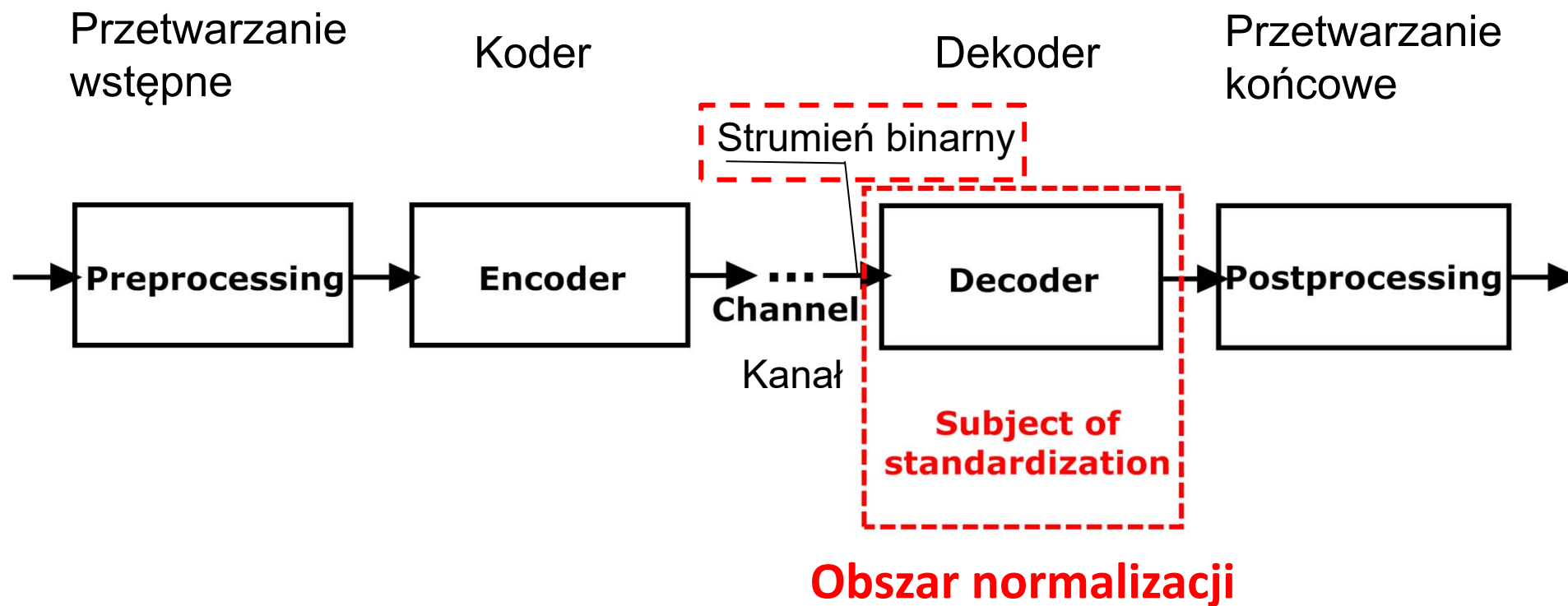
Umożliwienie współpracy produktom różnych wytwórców.

Na przykład dla kompresji:

Przedmiot normalizacji:

Semantyka i składnia strumienia danych

Przykład: normalizacja kompresji



Przykład normalizacji

Znormalizowany format pliku
Kontener
np. JFIF

Znormalizowany strumień binarny
np. JPEG

Zasady normalizacji

- Porozumienie
- Zapotrzebowanie przemysłu
- Dobrowolność

Normy są tworzone przez:

- organizacje międzynarodowe ,
- konsorcja firm .

Organizacje międzynarodowe

ISO – International Organization for Standardization, działa od 1947r.

Międzynarodowa Organizacja Normalizacyjna

normalizacja we wszystkich obszarach techniki i gospodarki

członkami są komitety narodowe (Polska jest pełnoprawnym członkiem reprezentowanym przez PKN – Polski Komitet Normalizacyjny)

IEC – International Electrotechnical Commission, działa od 1906 r.

Międzynarodowa Komisja Elektrotechniczna

normalizacja w zakresie elektrotechniki, elektroniki i pokrewnych obszarów techniki

członkami są komitety narodowe (Polska jest pełnoprawnym członkiem reprezentowanym przez PKN – Polski Komitet Normalizacyjny)

ITU – International Telecommunication Union, agencja ONZ działa od 1865 r. (początkowo jako Int. Telegraph Union),

Międzynarodowy Związek Telekomunikacyjny

normalizacja w zakresie telekomunikacji

członkami są kraje (Polska od 1921r.) oraz firmy/instytucje.

ITU-T - Telecommunication Standardization Sector,

ITU-R - Radiocommunication Sector.

Inne organizacje o znaczeniu międzynarodowym

ETSI – European Telecommunications Standards Institute
założony w 1988.

IEEE – The Institute of Electrical and Electronics Engineers
założony w 1884r. jako American Institute of Electrical Engineers
IEEE od 1963 z połączenia z IRE (Institute of Radio Engineers)

Instytut Inżynierów Elektryków i Elektroników

organizacja zawodowa zajmująca się również normalizacją
członkami są osoby fizyczne (w tym studenci)

SMPTE – Society of Motion Picture and Television Engineers,
działa od 1916 r.

Towarzystwo Inżynierów Filmowych i Telewizyjnych

normalizacja w zakresie kina i telewizji
członkami są osoby fizyczne i instytucje

Chiny

AVS – Audio Video Coding Standard (także grupa robocza, konsorcjum firm
i organizacja normalizacyjna)

Organizacja ISO

- **Technical Committee** – jest ich ponad 250.
Np. JTC1 – Joint (ISO+IEC) Technical Committee :
Information Technology.
- **Subcommittee** – podkomitet , np.: JTC1/SC29 – Coding of audio, picture, multimedia and hypermedia information.
- **Working group** – grupa ekspertów, praca zorganizowana w cyklicznych spotkaniach, np. w JTC1/SC29:
WG1 – JPEG Coding of digital representations of images,
WG2 – MPEG Technical requirements,
WG3 – MPEG Systems,
WG4 – MPEG Video coding,
WG5 – MPEG Joint Video Coding Team(s) with ITU-T SG 16 (JVET),
WG6 – MPEG Audio coding,
itd.
Objaśnienia: JPEG = Joint Photographic Experts Group
 MPEG = Moving Picture Experts Group
 JVET = Joint Video Experts Team

Tworzenie propozycji norm w grupach roboczych

- **Wymagania (Requirements)**
 - Wymagane zainteresowanie przemysłu
 - Wezwanie do ujawnienia dowodów istnienia techniki (Call for Evidence)
- **Faza konkurencji (Competitive phase)**

Wezwanie propozycji (Call for Proposals) ⇒ Testowanie propozycji ⇒ Wybór podstawowych metod
- **Faza współpracy (Collaborative phase)**

Poprawa technik, tworzenie opisu metod normatywnych i nienormatywnych oraz oprogramowania modelowego (Test Model)
Tworzenie roboczych wersji tekstu normy (Working Draft)
- **Faza komitetu** – Committee Draft – roboczy tekst normy oceniany przez komitet
- **Robocza norma międzynarodowa (Draft International Standard – DIS)**
- **Norma międzynarodowa (International Standard – IS)**

Międzynarodowe normy

- **ISO / IEC**

Norma międzynarodowa - International Standard (IS)

e.g.: ISO/IEC IS 14496 "Information technology

-- Coding of audio-visual objects"

-- Podzielona na części (arkusze), na przykład:

e.g. "Part 10: Advanced Video Coding" oznaczona jako ISO/IEC IS 14496-10: 2010 – oznacza wydanie z 2010 r.

- Konkretnie wydanie konkretnego arkusza może mieć:

- korekty (corrigenda),

- uzupełnienia (amendments).

- Potem są one scalane w kolejnym wydaniu:

np. ISO/IEC IS 14496-10: 2014 (które następnie może mieć nowe korekty i uzupełnienia).

Międzynarodowe normy

- **ITU-T**

Zalecenie - Recommendation (Rec.)

np. ITU-T Rec. H.264 "Advanced video coding for generic audiovisual services"

oznaczona jako ITU-T Rec. H.264: 2010

– oznacza wydanie z 2010 r.

NIE dzieli się na części.

- **Równoważność:** Na przykład:

ISO/IEC IS 14496-10 = ITU-T Rec. H.264

czyli część (arkusz) normy ISO = zalecenie ITU.

Numery i daty wydań nie zawsze zgodne.

1.5. Patenty

- Patent to ograniczone w czasie prawo właściciela do wyłącznego korzystania z wynalazku w celach zawodowych lub zarobkowych.
- Innymi słowy: to prawo do zabrania innym do komercyjnego korzystania z pewnego rozwiązania.
- Prawo obowiązuje na terenie konkretnego kraju we wskazanym okresie (zwykle do 20 lat).
- Prawo uzyskiwane na podstawie decyzji administracyjnej – wydaje ją urząd patentowy danego kraju.

Urząd Patentowy RP - UPRP

United States Patent and Trademark Office – USPTO

W USA patenty wydawane od 1790 r.,
1810 założenie urzędu.

Międzynarodowe instytucje patentowe

- **Europejski Urząd Patentowy**
European Patent Office EPO
(Monachium, Haga, Wiedeń, Berlin)
- **Światowa Organizacja Własności Intelektualnej**
World Intellectual Property Organization WIPO

Zakres merytoryczny patentów

Wartość techniczna – różnie rozumiana

Problem programów komputerowych i algorytmów

Procedura uzyskiwania patentu

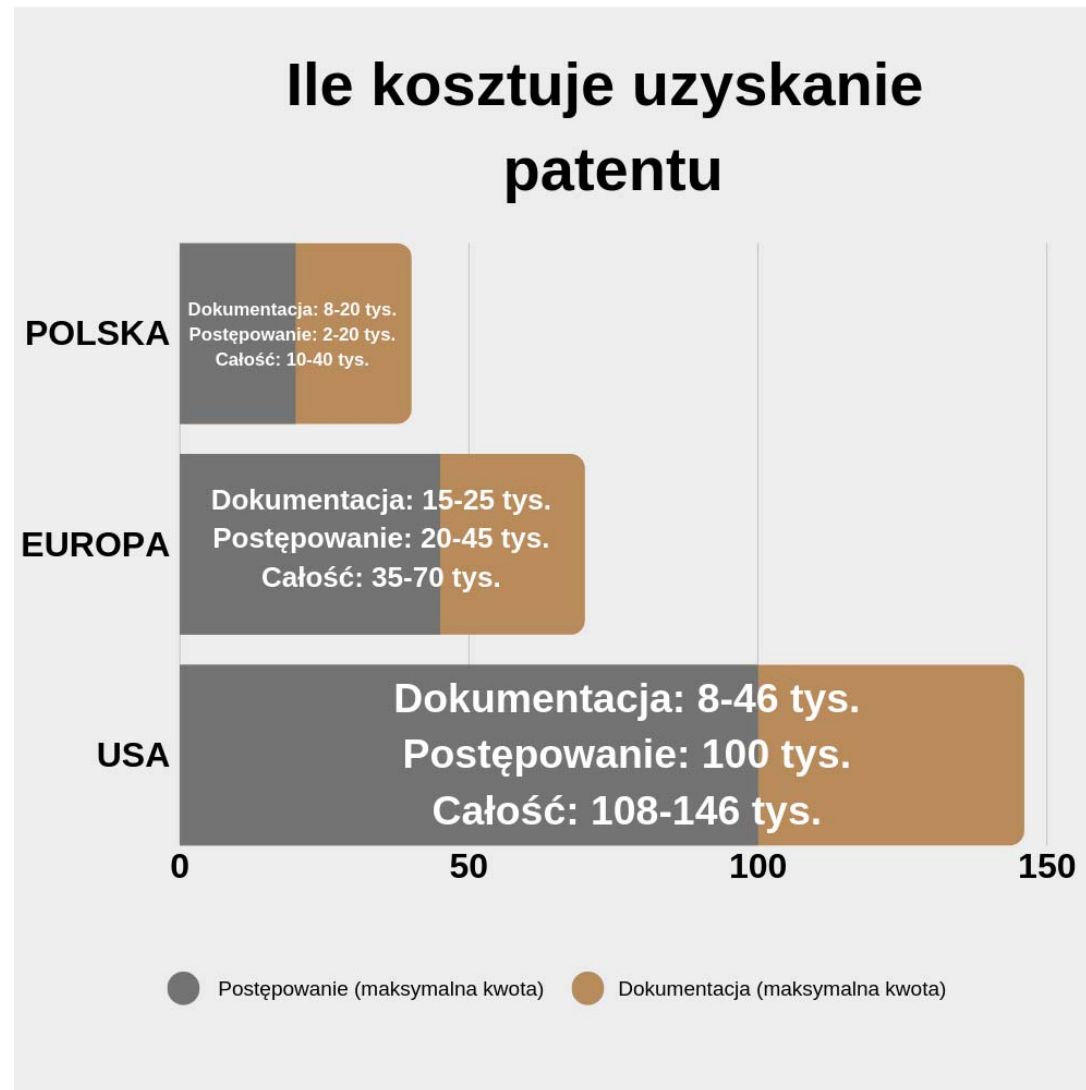
- **Zgłoszenie patentowe**

Zgłaszane do urzędu patentowego

Zgłoszenie może zrealizować tylko rzecznik patentowy

- Z zasady zgłoszenie patentowe nie jest ujawniane przez pewien okres, np. 18 miesięcy po czym jest publikowane.
- Zgłoszenie jest oceniane przez pracowników urzędu patentowego.
- Jeżeli zostanie zaopiniowane pozytywnie, to należy wnieść opłatę za ochronę na pewien okres. Potem możliwe są przedłużenia zwykle do 20 lat.

Koszty uzyskania patentu



Patenty w multimedialach

- Szczególnie wartościowe są patenty odnoszące się do rozwiązań ujętych w normach.
- Niektóre normy są związane z tysiącami patentów.
- Patenty dla firmy są inwestycją, która pozwala na utrzymanie pozycji rynkowej.
- Instytuty i laboratoria badawcze sprzedają pule patentów i w ten sposób częściowo finansują swoją działalność.

Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego na studiach na kierunkach Elektronika i Telekomunikacja oraz Teleinformatyka.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

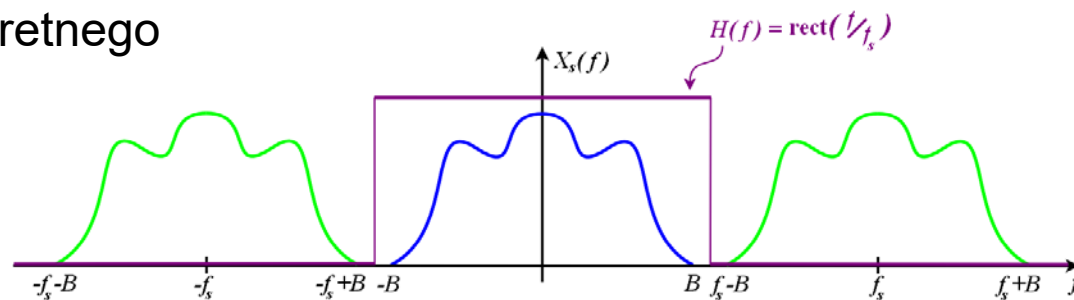
2. Akwizycja, reprezentacja i percepcja obrazów

2.1. Próbkowanie obrazu w kamerze

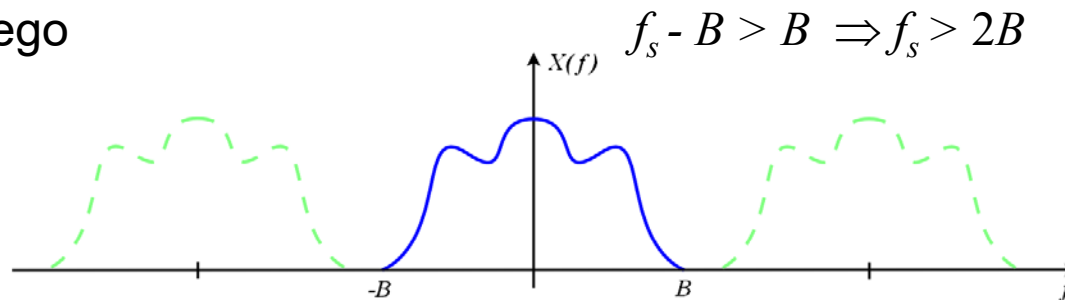
Twierdzenie o próbkowaniu

Aby odtworzyć sygnał analogowy z próbek, należy próbkować sygnał analogowy z częstotliwością co najmniej dwukrotnie większą niż największa częstotliwość w widmie sygnału analogowego.

Widmo sygnału dyskretnego



Widmo sygnału ciągłego



Sygnał o ograniczonym paśmie

Akwizycja fonii



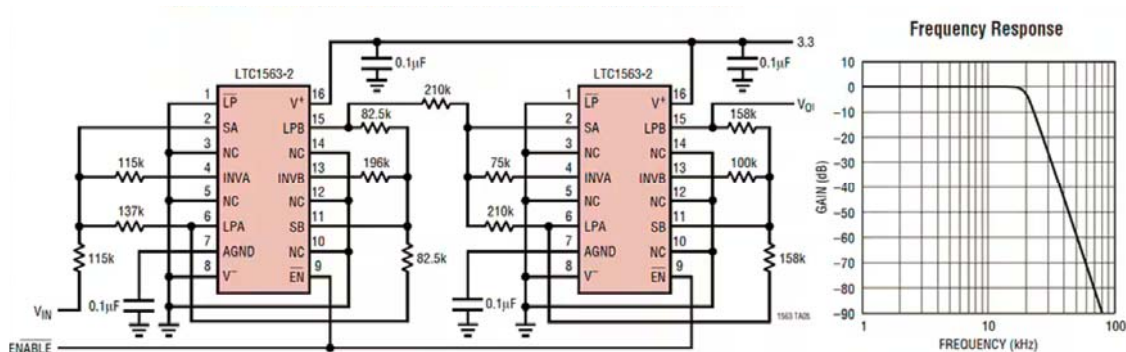
Sygnał ciągły analogowy



$f(t)$

Mikrofon
= Przetwornik
akustyczno-
-elektryczny

Analogowy filtr dolnoprzepustowy



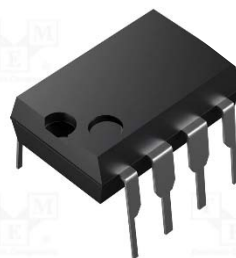
Sygnał ciągły analogowy



Układ
próbujący
z pamięcią



Sygnał
analogowy
schodkowy



Przetwornik
analogowo - cyfrowy

Sygnał
cyfrowy

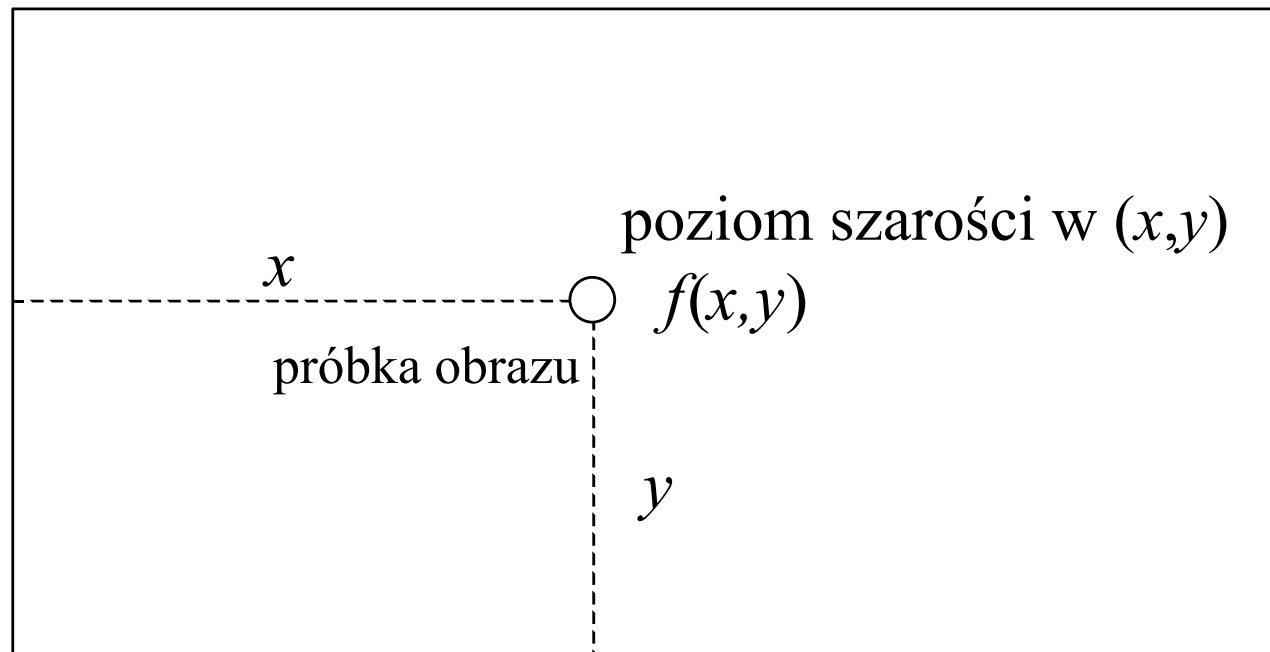


Obraz – sygnał dwuwymiarowy

Sygnał dwuwymiarowy (*2-dimensional signal*) jest opisany funkcją **dwóch** zmiennych niezależnych.

$$f(x, y)$$

OBRAZ

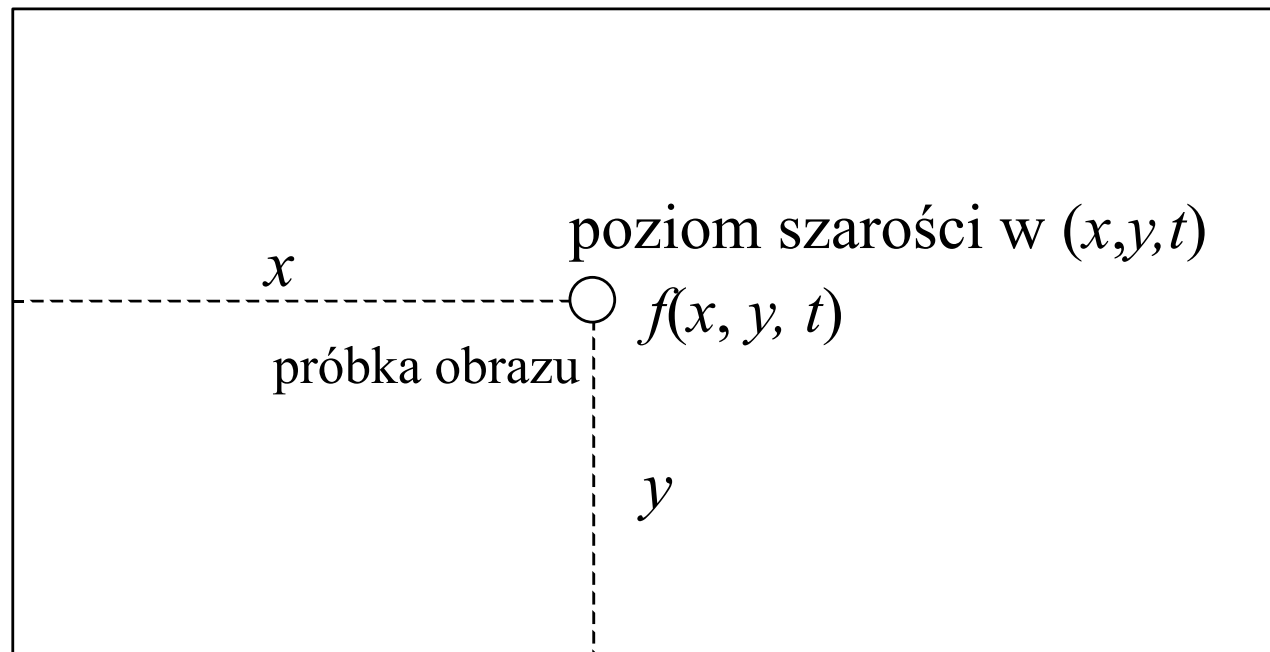


Wizja – sygnał trójwymiarowy

Sygnał trójwymiarowy (*3-dimensional signal*) jest opisany funkcją **trzech** zmiennych niezależnych.

$$f(x, y, t)$$

OBRAZ w chwili t



Częstotliwości przestrzenne



Okres jest mierzony
w jednostkach długości

T_1

$$f_1 = \frac{1}{T_1} \quad \text{oraz} \quad \omega_1 = \frac{2\pi}{T_1} = 2\pi f_1$$

Częstotliwość: Jednostki: 1/mm, 1/cm, 1/inch

Pulsacja: Jednostki: rad/mm, rad/cm, rad/inch .

Częstotliwości przestrzenne mierzone w kierunku pionowym



$$\omega_2 = \frac{2\pi}{T_2} = 2\pi f_2, \text{ gdzie } f_2 = \frac{1}{T_2}.$$

Częstotliwości sygnału dwuwymiarowego



1 cm

$$f_1 = 2 \text{ [1/cm]} \quad \omega_1 = 2\pi \cdot 2 \text{ [rad/cm]} = 4\pi \text{ rad/cm}$$

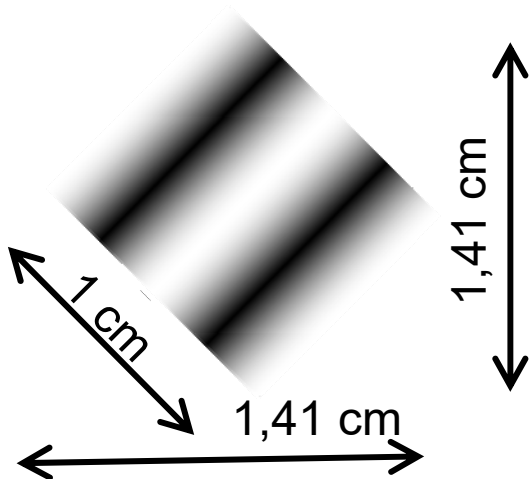
$$f_2 = 0 \text{ [1/cm]} \quad \omega_2 = 0 \text{ [rad/cm]}$$



1 cm

$$f_1 = 0 \text{ [1/cm]}$$

$$f_2 = 2 \text{ [1/cm]}$$



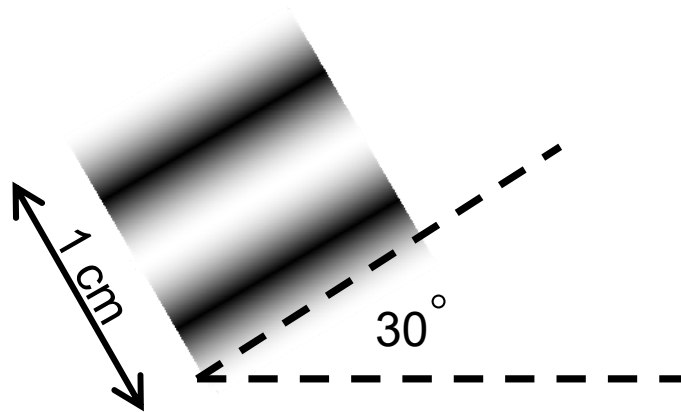
1,41 cm

$$f_1 = 2/1,41 = 1,41 \text{ [1/cm]}$$

$$f_2 = 2/1,41 = 1,41 \text{ [1/cm]}$$

Przykład

- Wyznaczyć częstotliwości sygnału z rysunku:



Wyznaczanie widma

Jednowymiarowe przekształcenie Fouriera dyskretnego sygnału

Transformata Fouriera (widmo sygnału)

$$F(j\omega) = \sum_{n=-\infty}^{\infty} f(n) \cdot e^{-jn\omega}$$

Przekształcenie odwrotne

$$f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(j\omega) \cdot e^{+jn\omega} d\omega$$

Dwuwymiarowe przekształcenie Fouriera dyskretnego sygnału

Widmo obrazu nieruchomego jest określone transformatą dwuwymiarową

$$F(j\omega_1, j\omega_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f(n_1, n_2) \cdot \exp[-j(n_1\omega_1 + n_2\omega_2)]$$

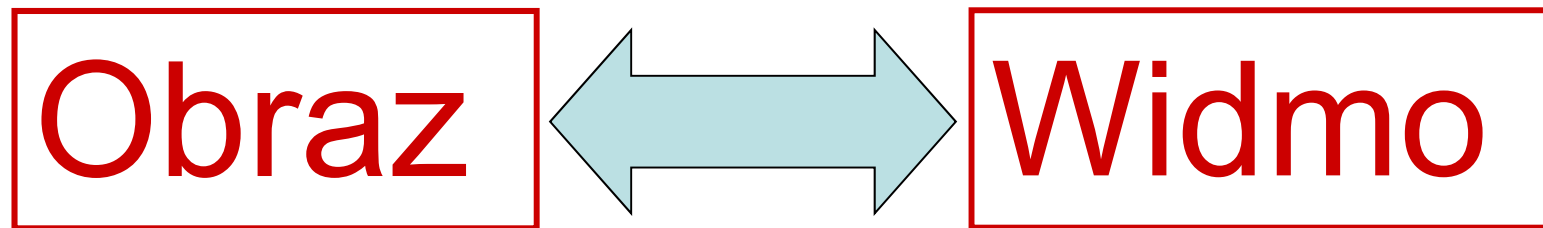
Przekształcenie odwrotne

$$f(n_1, n_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(j\omega_1, j\omega_2) \cdot \exp[+j(n_1\omega_1 + n_2\omega_2)] d\omega_1 d\omega_2$$

Dwie reprezentacje obrazu

Za pomocą próbek obrazu

Za pomocą próbek widma



$$f(n_1, n_2)$$

$$\leftrightarrow$$

$$F(j\omega_1, j\omega_2)$$

Funkcja położenia

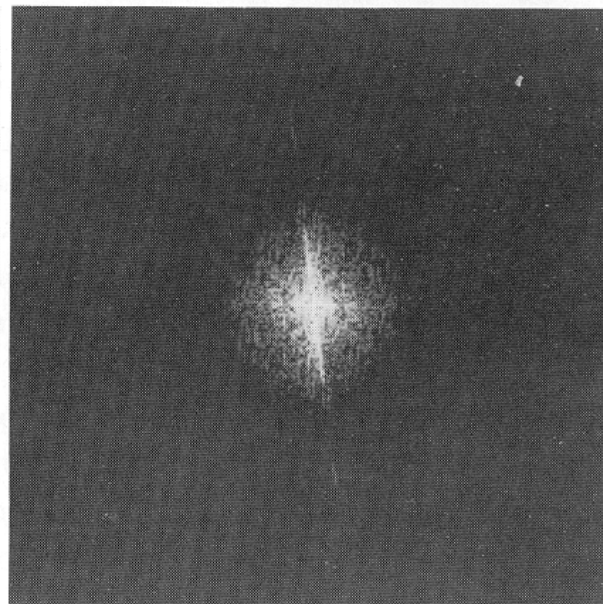
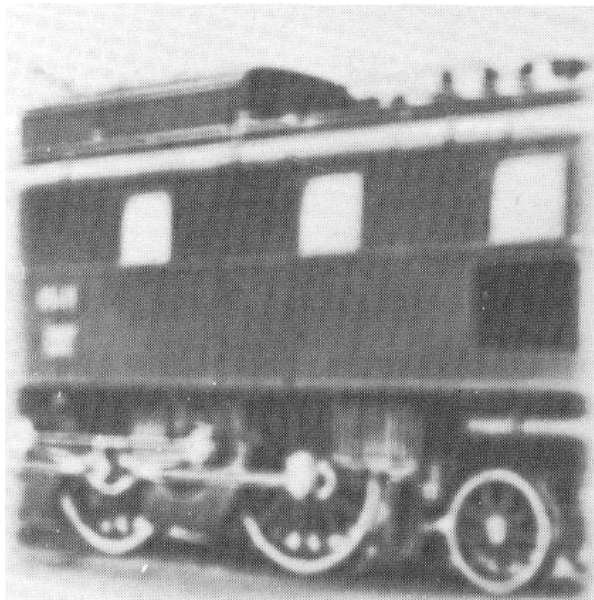
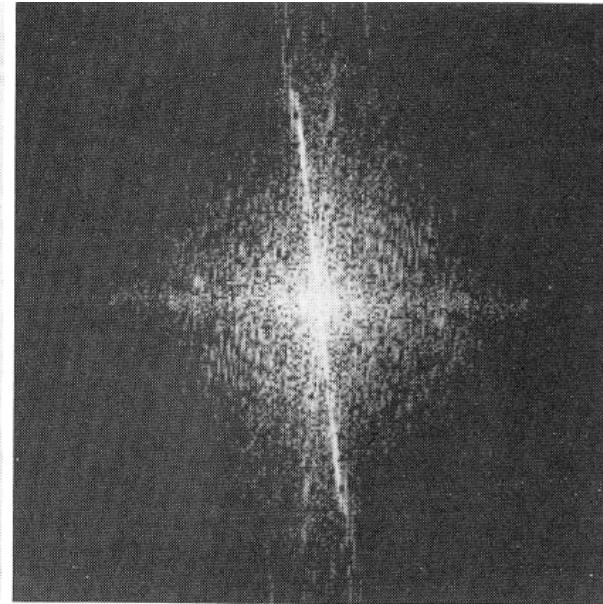
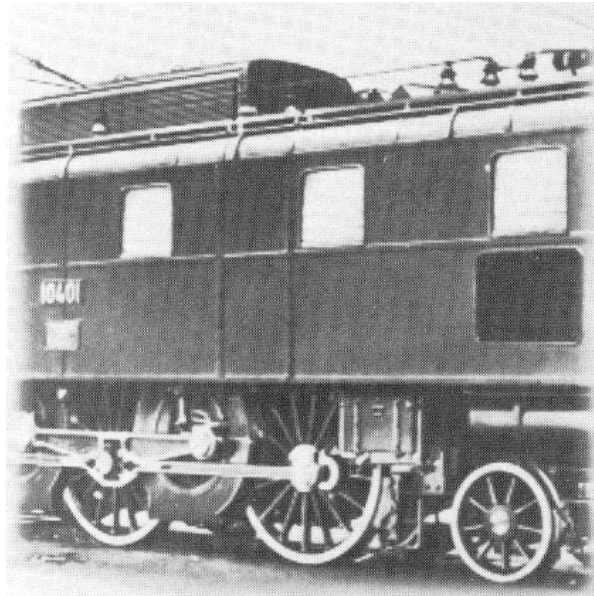
Funkcja częstotliwości

2 zmienne

$$n_1, n_2$$

$$\omega_1, \omega_2 (f_1, f_2)$$

Widma amplitudowe obrazu

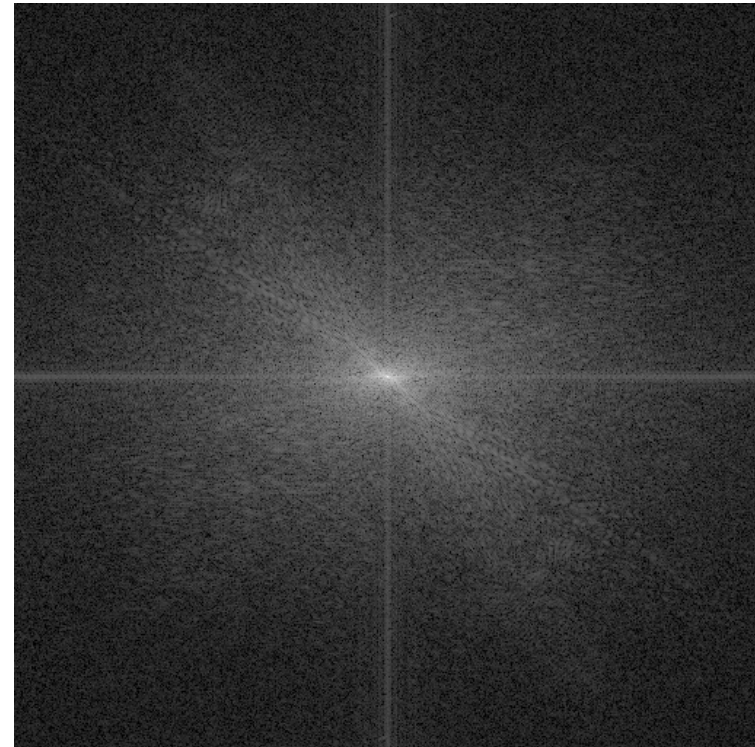


Widma amplitudowe obrazu

Obraz



Widmo amplitudowe



Obraz odtworzony z samego widma amplitudowego

Obraz oryginalny $f(n_1, n_2)$



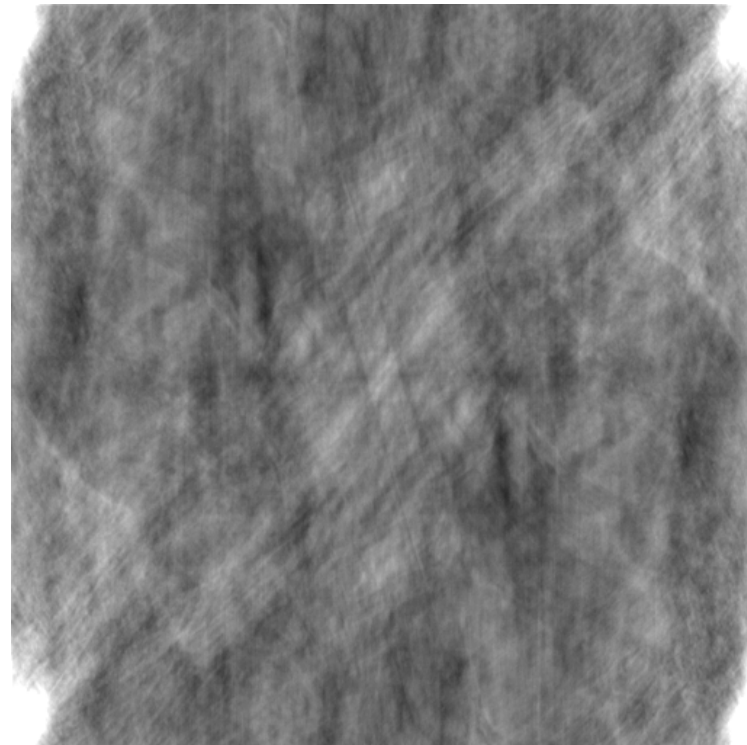
Transformata

$$F = |F(\omega_1, \omega_2)|e^{-j\varphi(\omega_1, \omega_2)}$$

Obraz odtworzony z widma amplitudowego

$$F_M = |F(\omega_1, \omega_2)|e^{-j0}$$

$$f_M = \mathcal{F}^{-1}[F_M]$$



Obraz odtworzony z samego widma fazowego

Obraz oryginalny $f(n_1, n_2)$



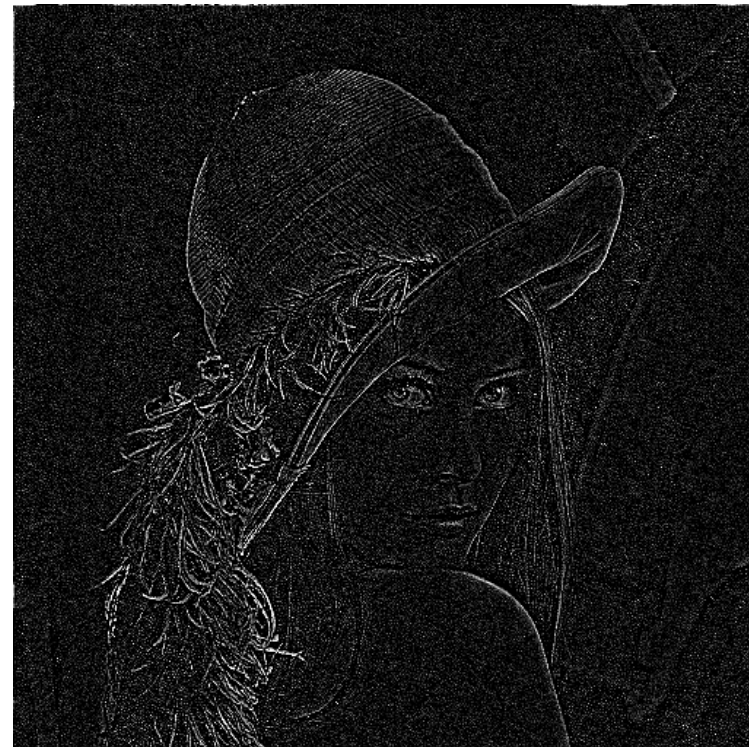
Transformata

$$F = |F(\omega_1, \omega_2)|e^{-j\varphi(\omega_1, \omega_2)}$$

Obraz odtworzony z widma amplitudowego

$$F_F = 1e^{-j\varphi(\omega_1, \omega_2)}$$

$$f_F = \mathcal{F}^{-1}[F_F]$$



Odtwarzanie obrazu z jego widma fazowego



Transformata

$$F_1 = \\ = |F_1(\omega_1, \omega_2)|e^{-j\varphi_1(\omega_1, \omega_2)}$$



Transformata

$$F_2 = \\ = |F_2(\omega_1, \omega_2)|e^{-j\varphi_2(\omega_1, \omega_2)}$$

Widmo fazowe „Lena”

Widmo amplitudowe „Boats”

$$f_L = \mathcal{F}^{-1}[|F_2(\omega_1, \omega_2)|e^{-j\varphi_1(\omega_1, \omega_2)}]$$



Odtwarzanie obrazu z jego widma fazowego



Transformata

$$F_1 = \\ = |F_1(\omega_1, \omega_2)|e^{-j\varphi_1(\omega_1, \omega_2)}$$



Transformata

$$F_2 = \\ = |F_2(\omega_1, \omega_2)|e^{-j\varphi_2(\omega_1, \omega_2)}$$

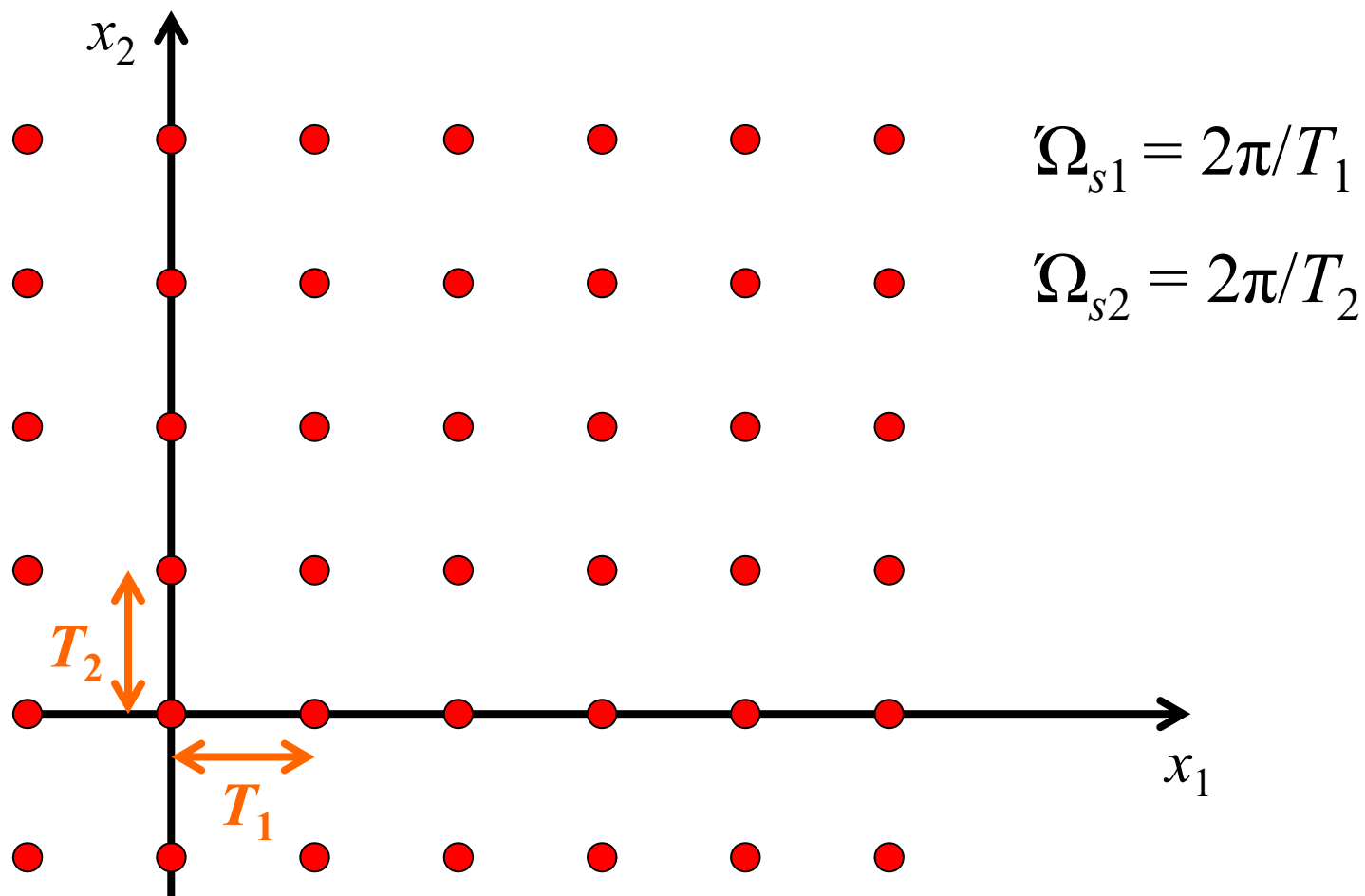
Widmo fazowe „Boats”

Widmo amplitudowe „Lena”

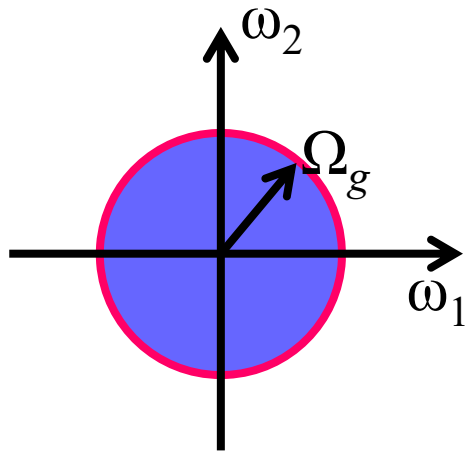
$$f_L = \mathcal{F}^{-1}[|F_1(\omega_1, \omega_2)|e^{-j\varphi_2(\omega_1, \omega_2)}]$$



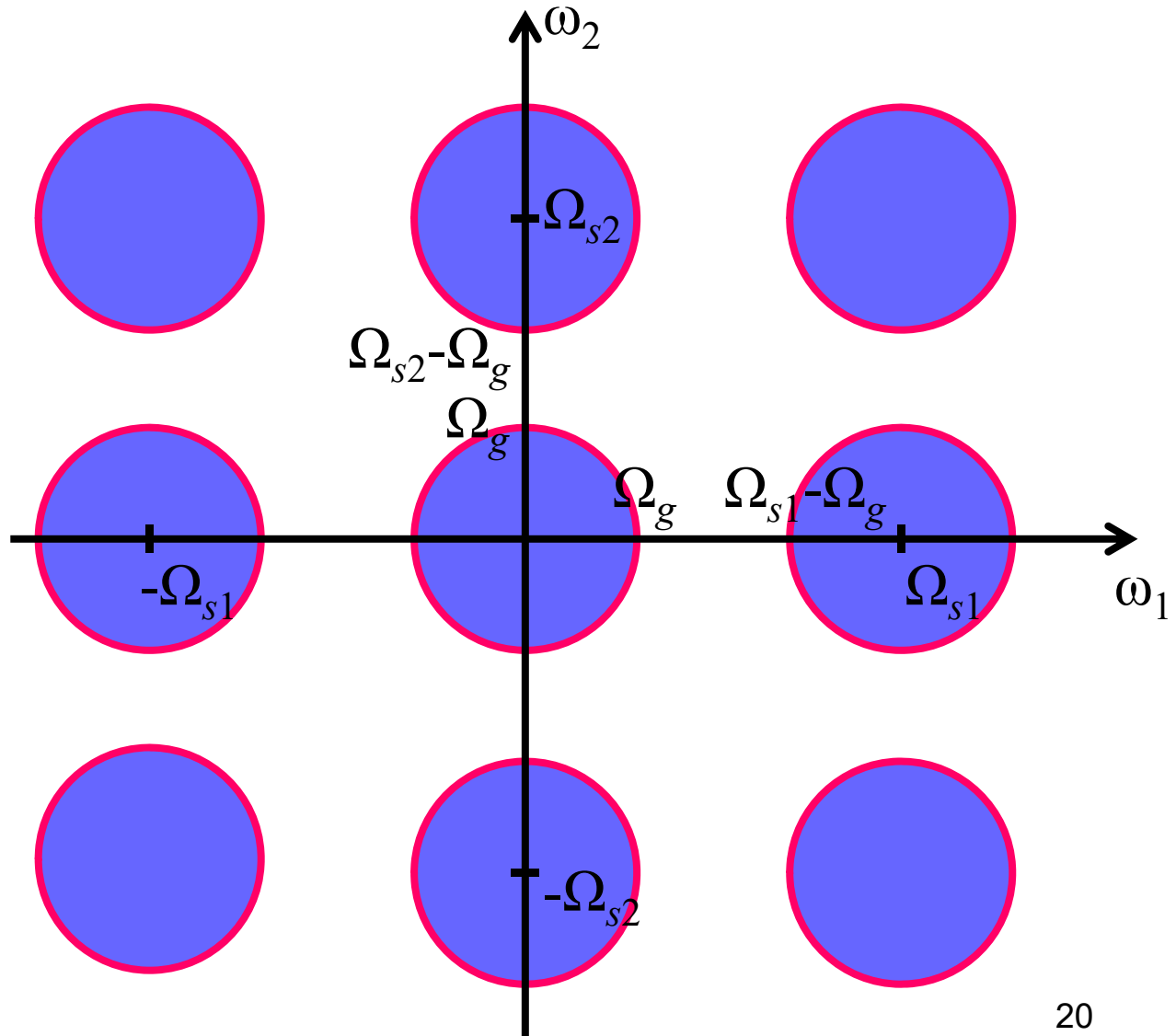
Próbkowanie sygnałów wielowymiarowych



Próbkowanie sygnałów wielowymiarowych

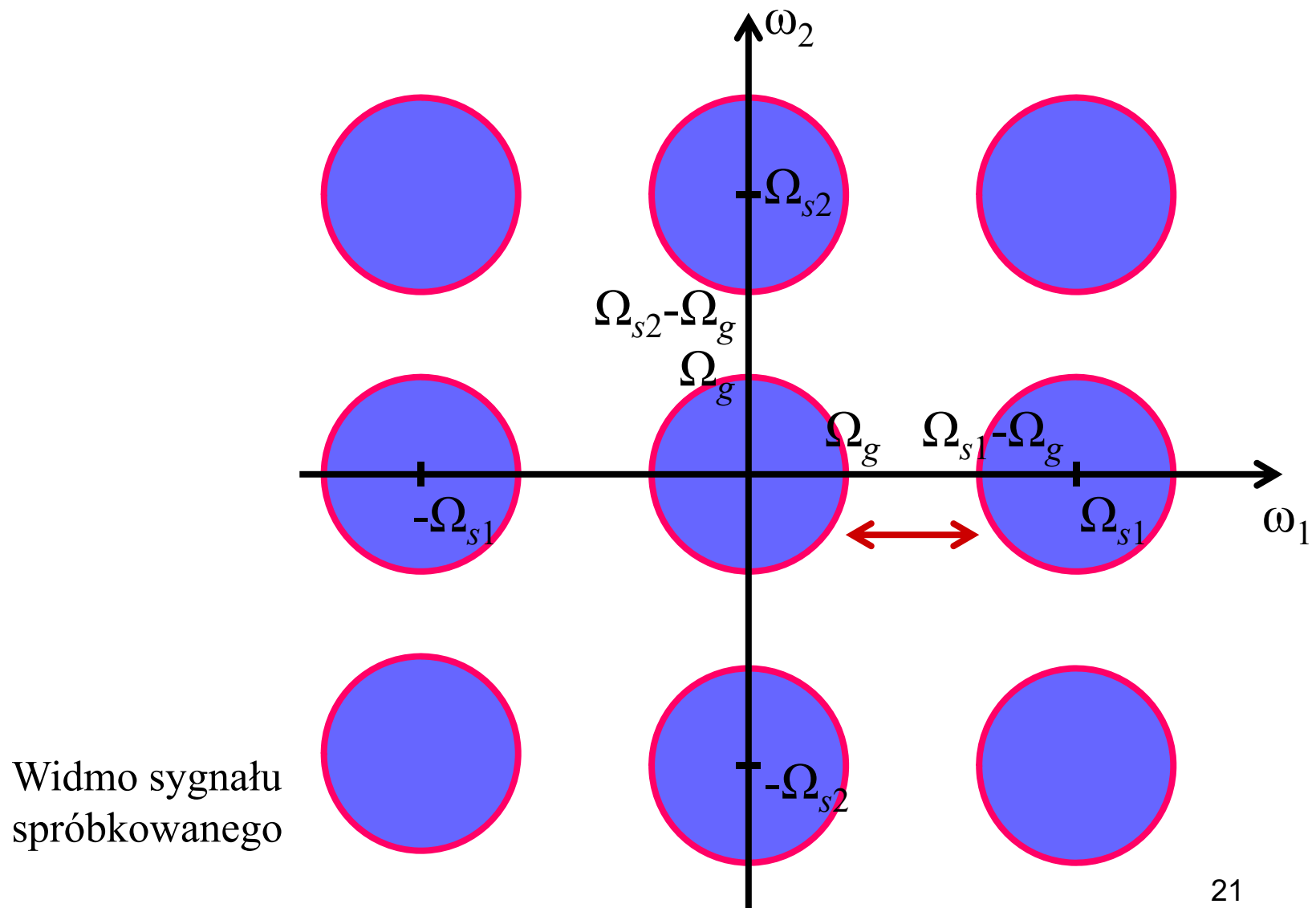


Widmo sygnału ciągłego
o ograniczonym paśmie



Widmo sygnału
spróbkowanego

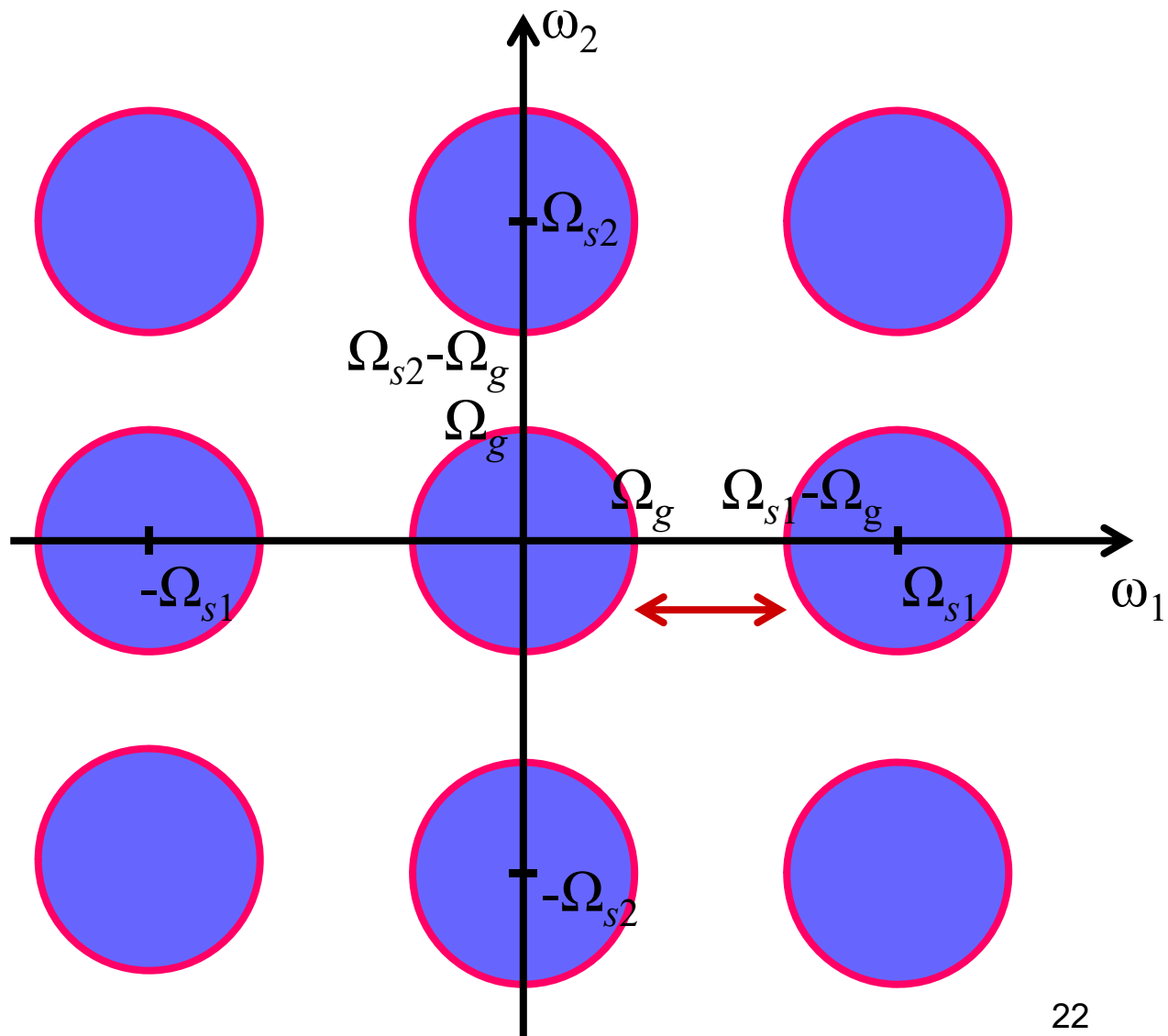
Twierdzenie o próbkowaniu



Twierdzenie o próbkowaniu

$$\Omega_{s1} - \Omega_g > \Omega_g$$

Widmo sygnału
spróbkowanego



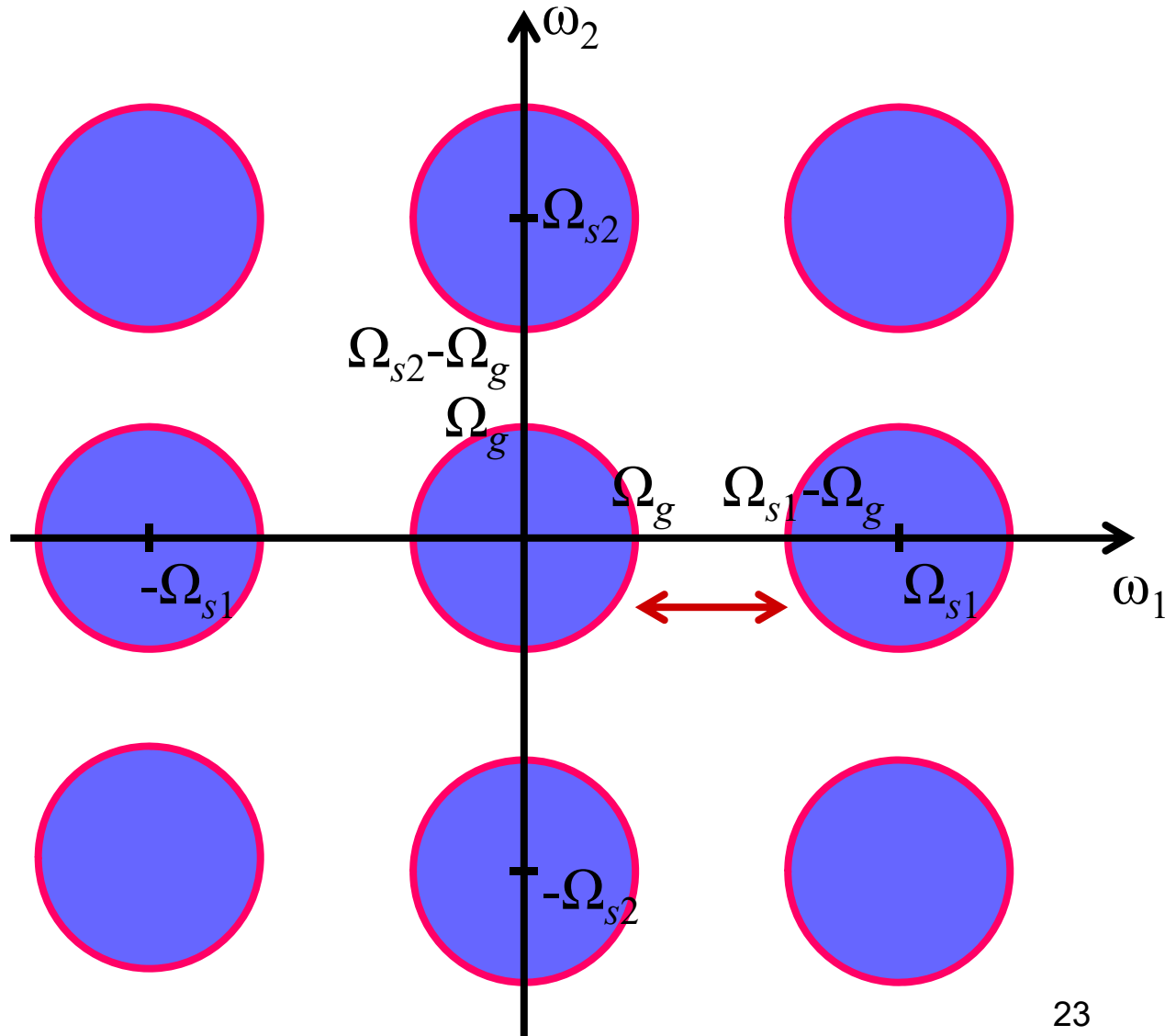
Twierdzenie o próbkowaniu

$$\Omega_{s1} - \Omega_g > \Omega_g$$



$$\Omega_{s1} > 2 \cdot \Omega_g$$

Widmo sygnału
spróbkowanego



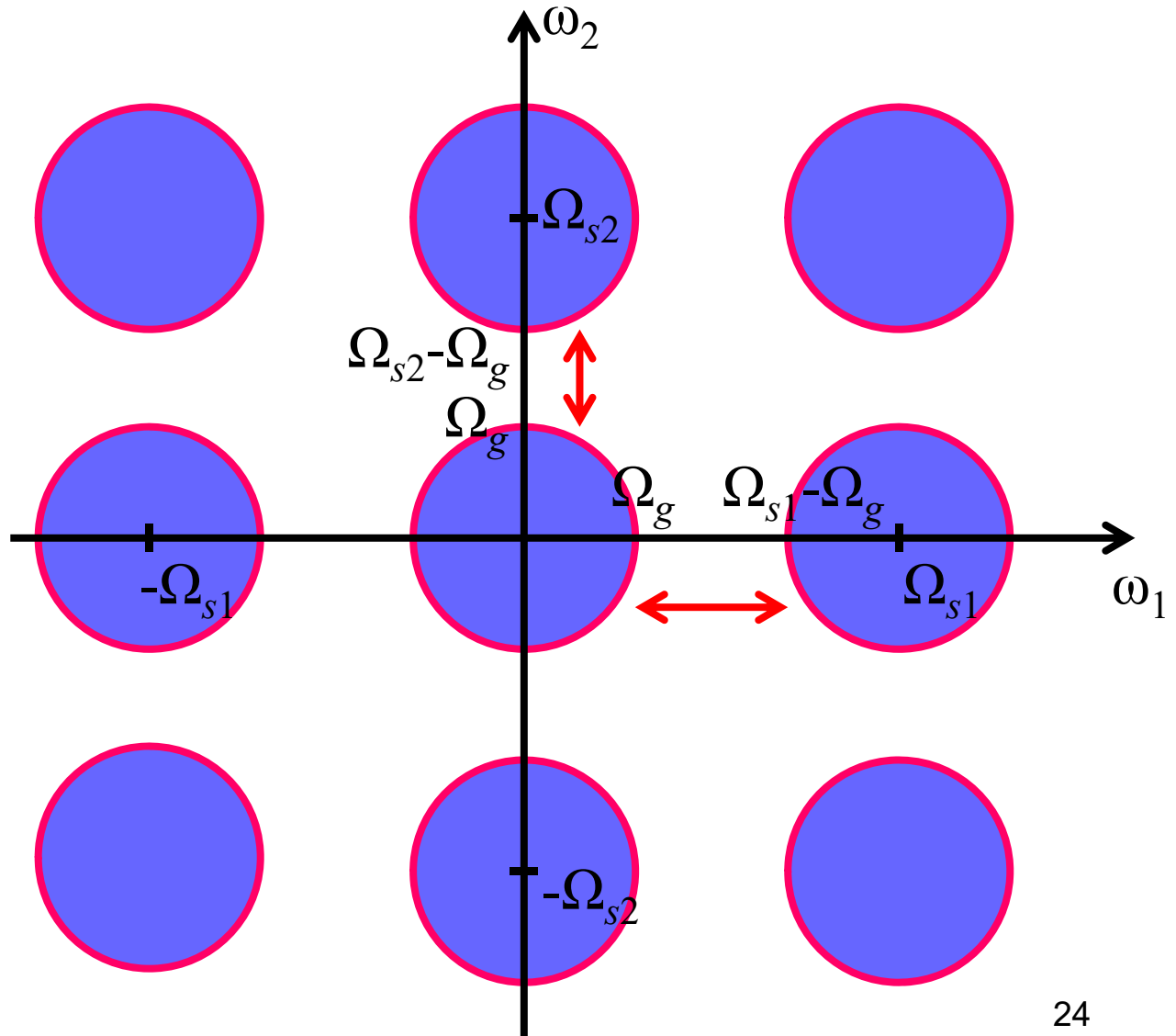
Twierdzenie o próbkowaniu

$$\Omega_{s1} - \Omega_g > \Omega_g$$



$$\Omega_{s1} > 2 \cdot \Omega_g$$

$$\Omega_{s2} - \Omega_g > \Omega_g$$



Twierdzenie o próbkowaniu

$$\Omega_{s1} - \Omega_g > \Omega_g$$

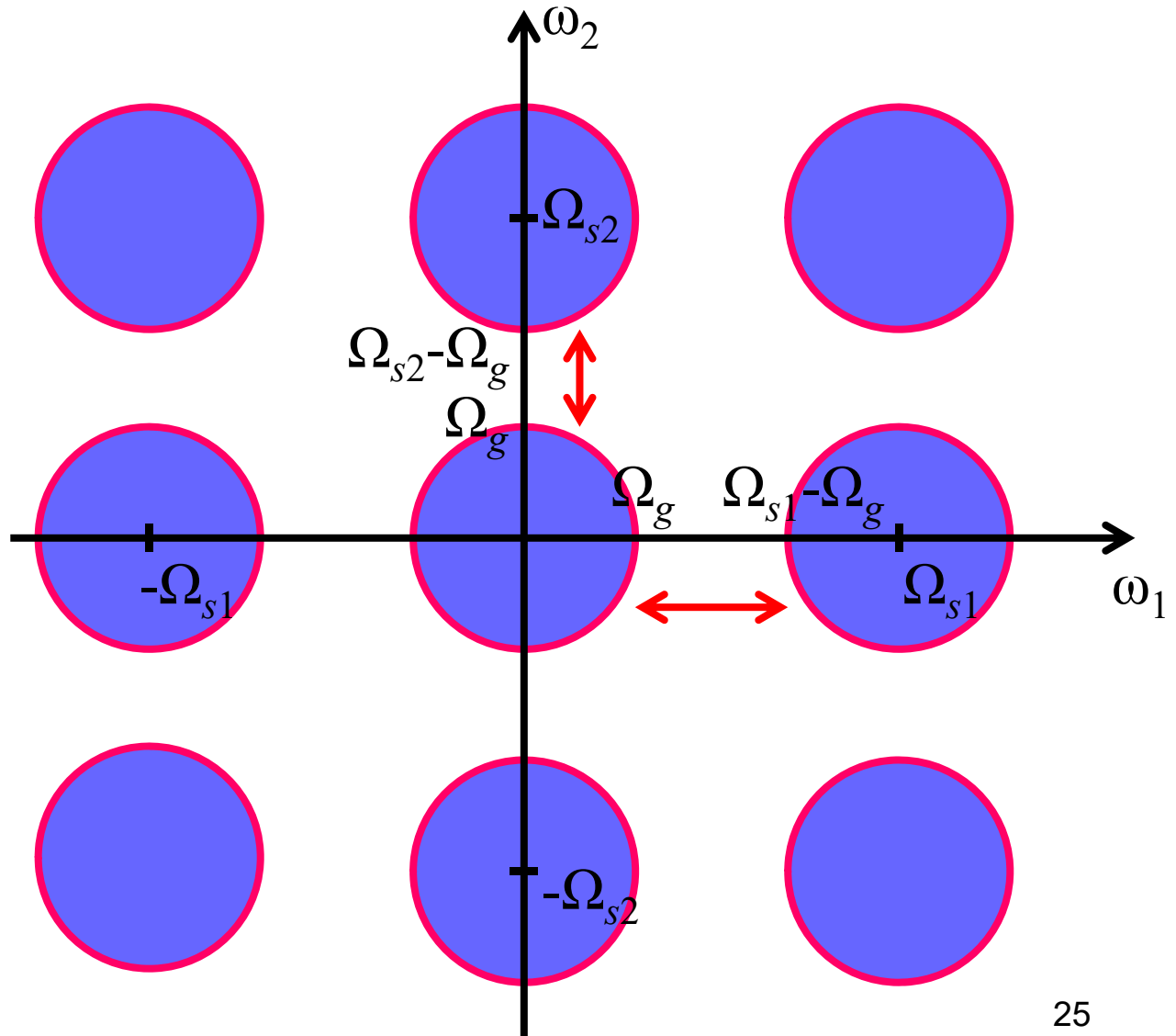


$$\Omega_{s1} > 2 \cdot \Omega_g$$

$$\Omega_{s2} - \Omega_g > \Omega_g$$



$$\Omega_{s2} > 2 \cdot \Omega_g$$



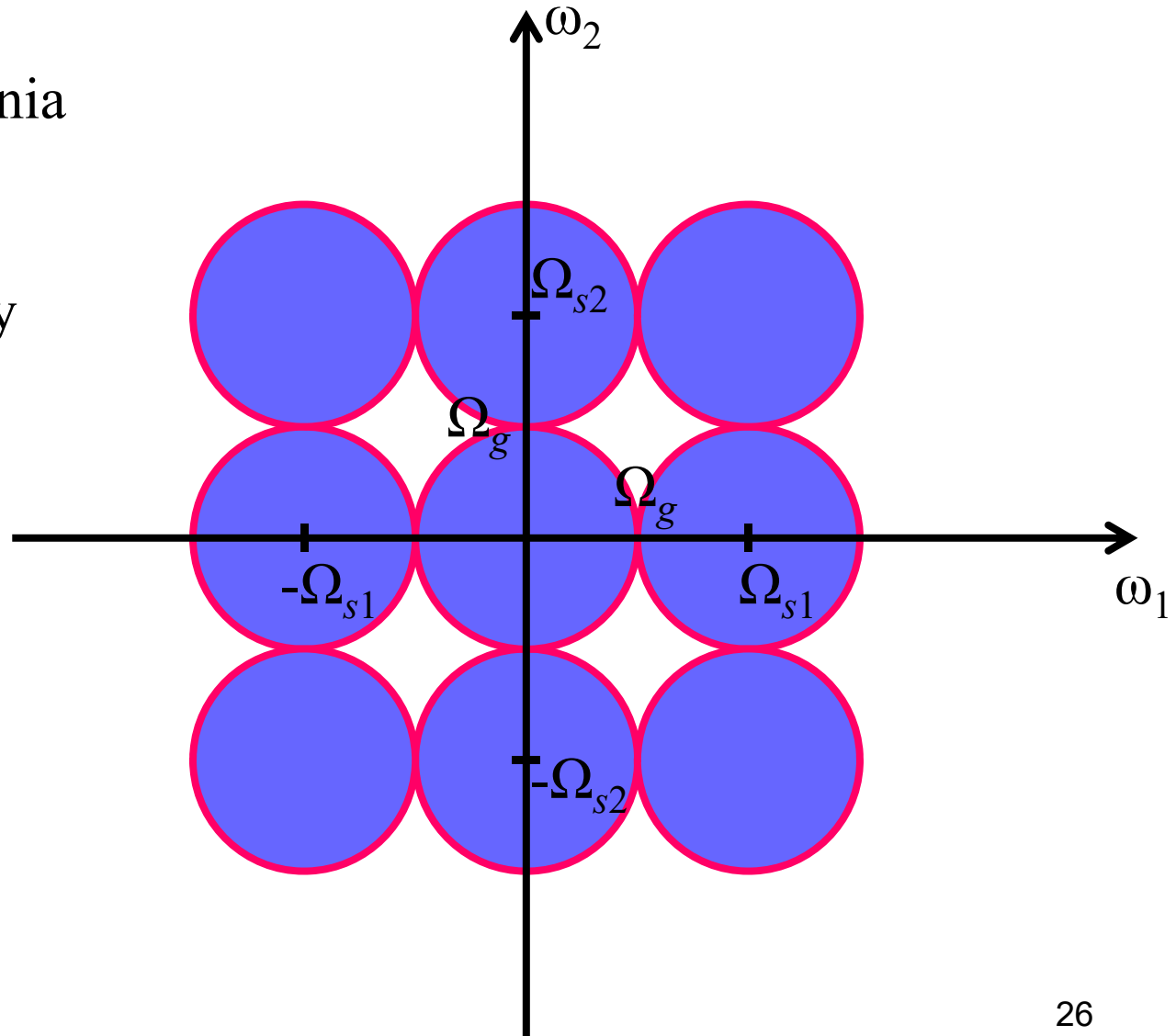
Twierdzenie o próbkowaniu

Zmniejszamy
gęstość próbkowania

czyli
wydłużamy okresy
próbkowania

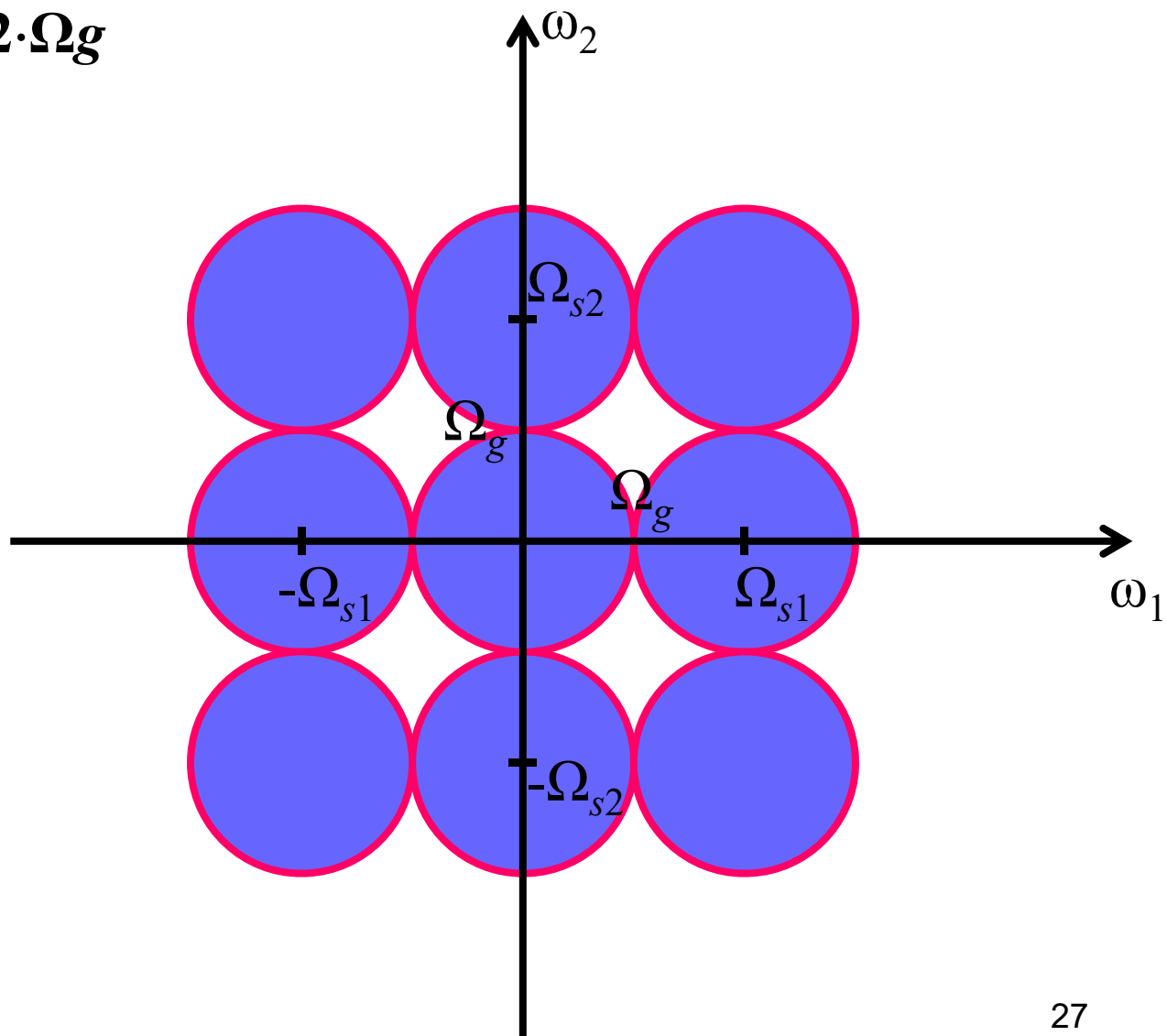
Przypadek
graniczny:

$$\Omega_{s1} = \Omega_{s2} = 2 \cdot \Omega_g$$



Przypadek graniczny

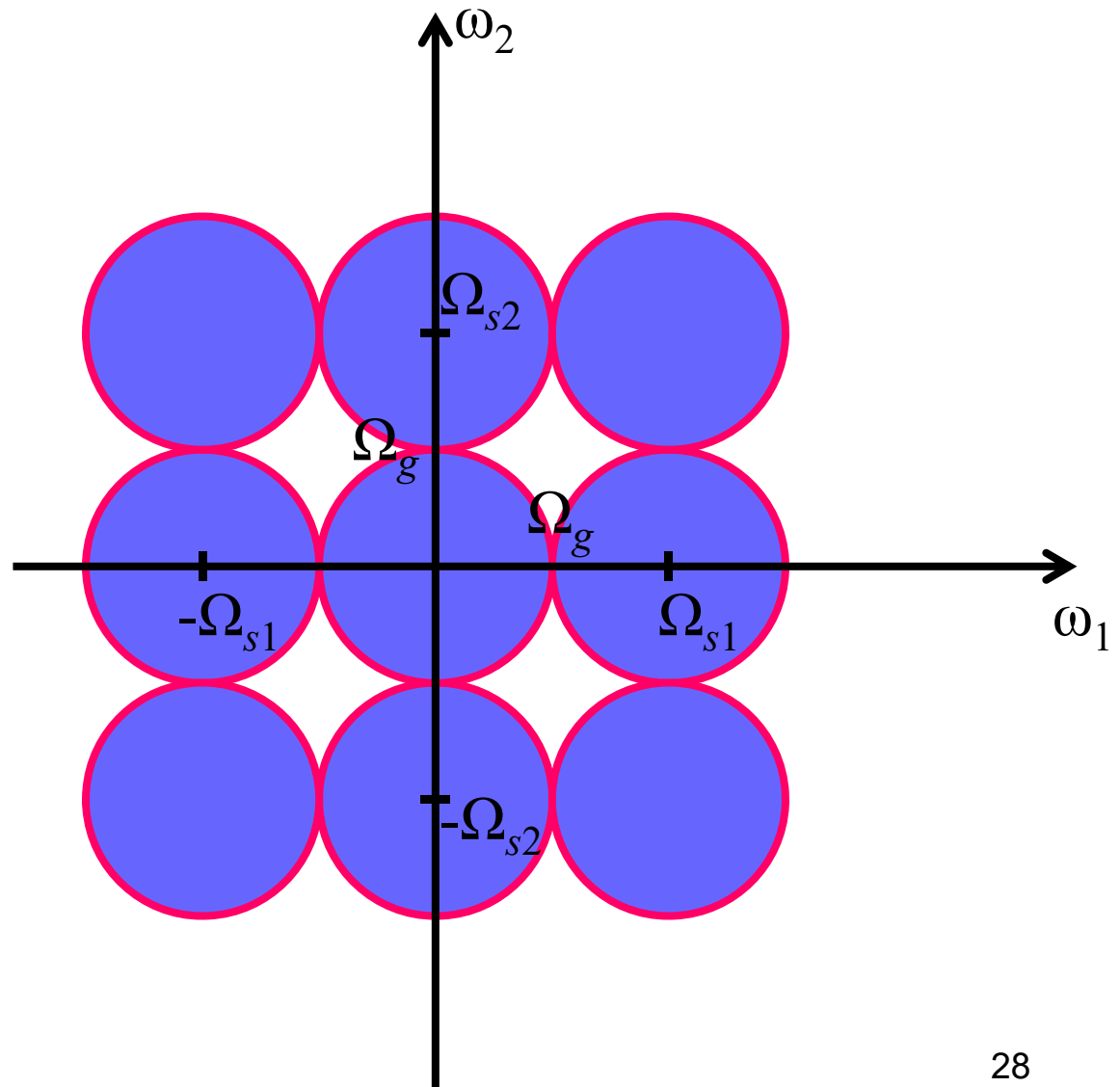
$$\Omega_{s1} = \Omega_{s2} = \Omega_s = 2 \cdot \Omega_g$$



Przypadek graniczny

$$\Omega_{s1} = \Omega_{s2} = \Omega_s = 2 \cdot \Omega_g$$

$$T_1 = T_2 = T$$

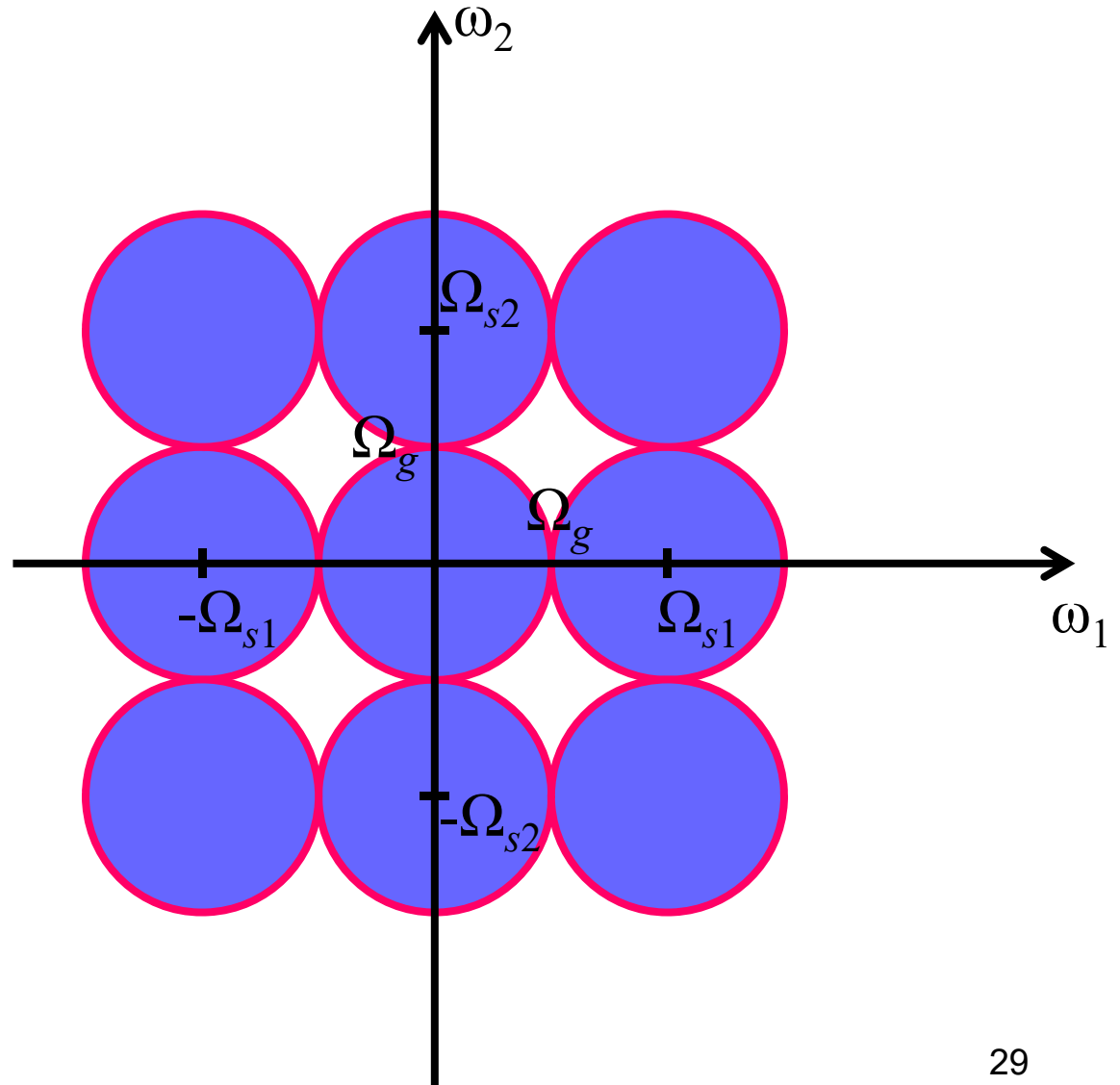


Przypadek graniczny

$$\Omega_{s1} = \Omega_{s2} = \Omega_s = 2 \cdot \Omega_g$$

$$T_1 = T_2 = T$$

$$T = 2\pi / \Omega_s = \pi / \Omega_g$$



Przypadek graniczny

$$\Omega_{s1} = \Omega_{s2} = \Omega_s = 2 \cdot \Omega_g$$

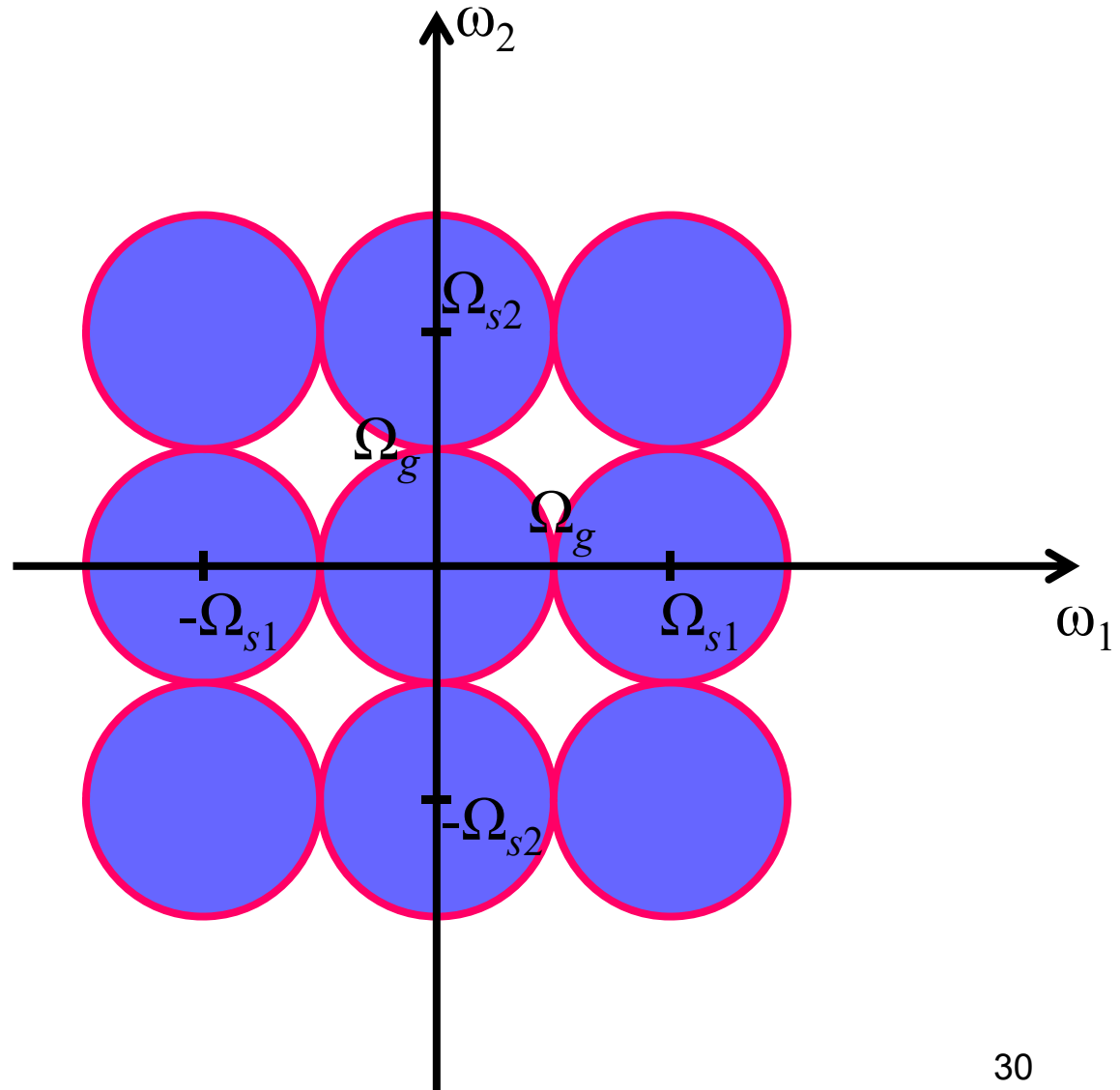
$$T_1 = T_2 = T$$

$$T = 2\pi / \Omega_s = \pi / \Omega_g$$

Minimalna

gęstość próbkowania:

$$1 / T^2 = \Omega_g^2 / \pi^2$$



Zniekształcenia obrazów wywołane zbyt małymi częstotliwościami próbkowania

Obrazy odtworzone z próbek wytworzonych:

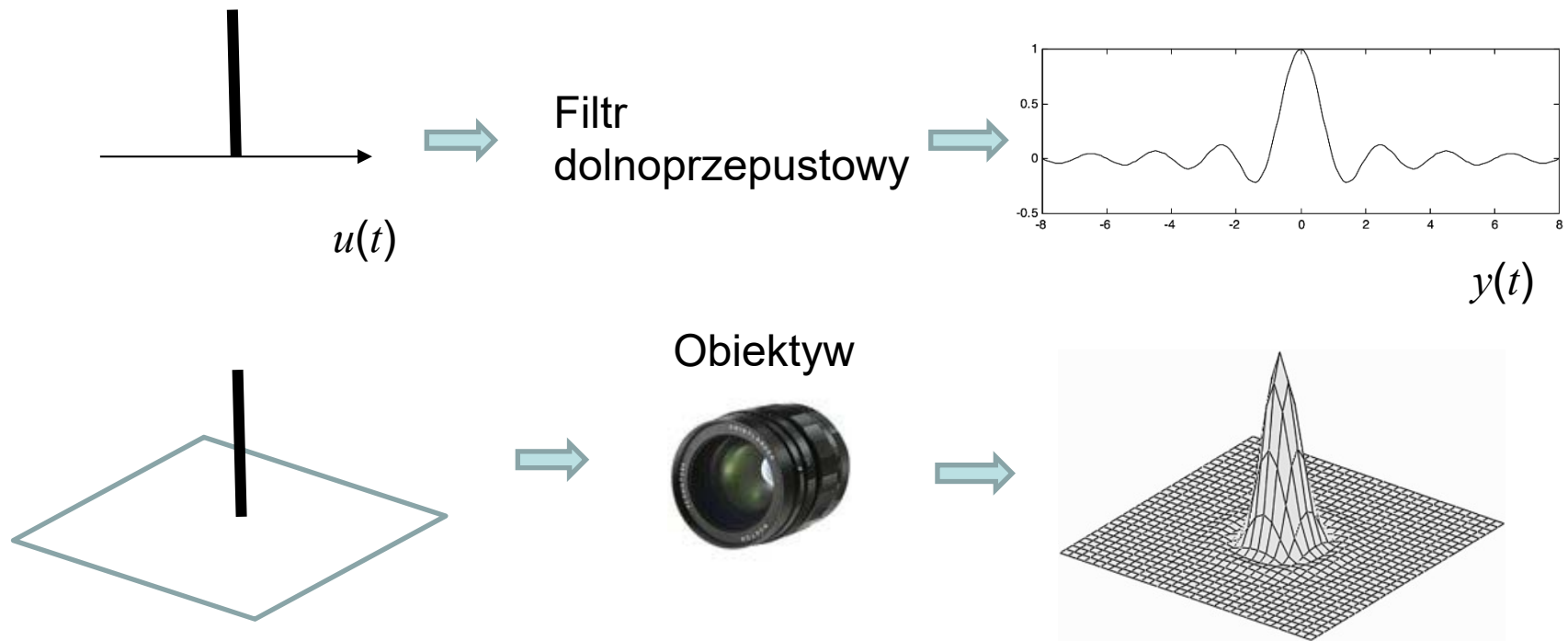
z wystarczającymi częstotliwościami próbkowania



z niewystarczającymi częstotliwościami próbkowania



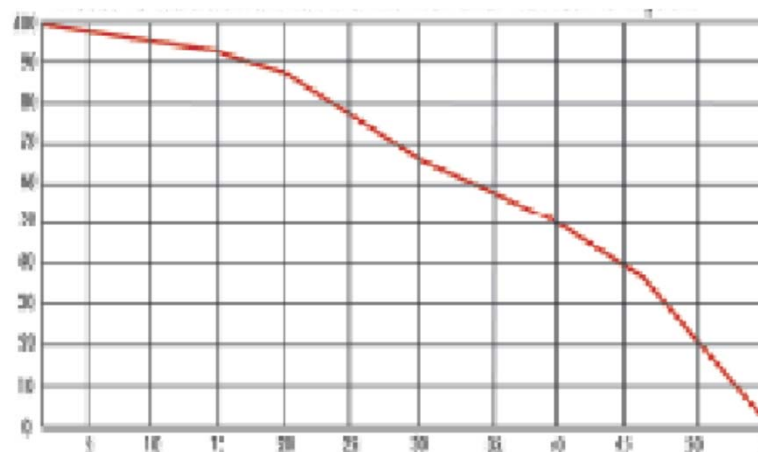
Filtracja dolnoprzepustowa w dziedzinie przestrzennej



Obiektyw jako filtr dolnoprzepustowy

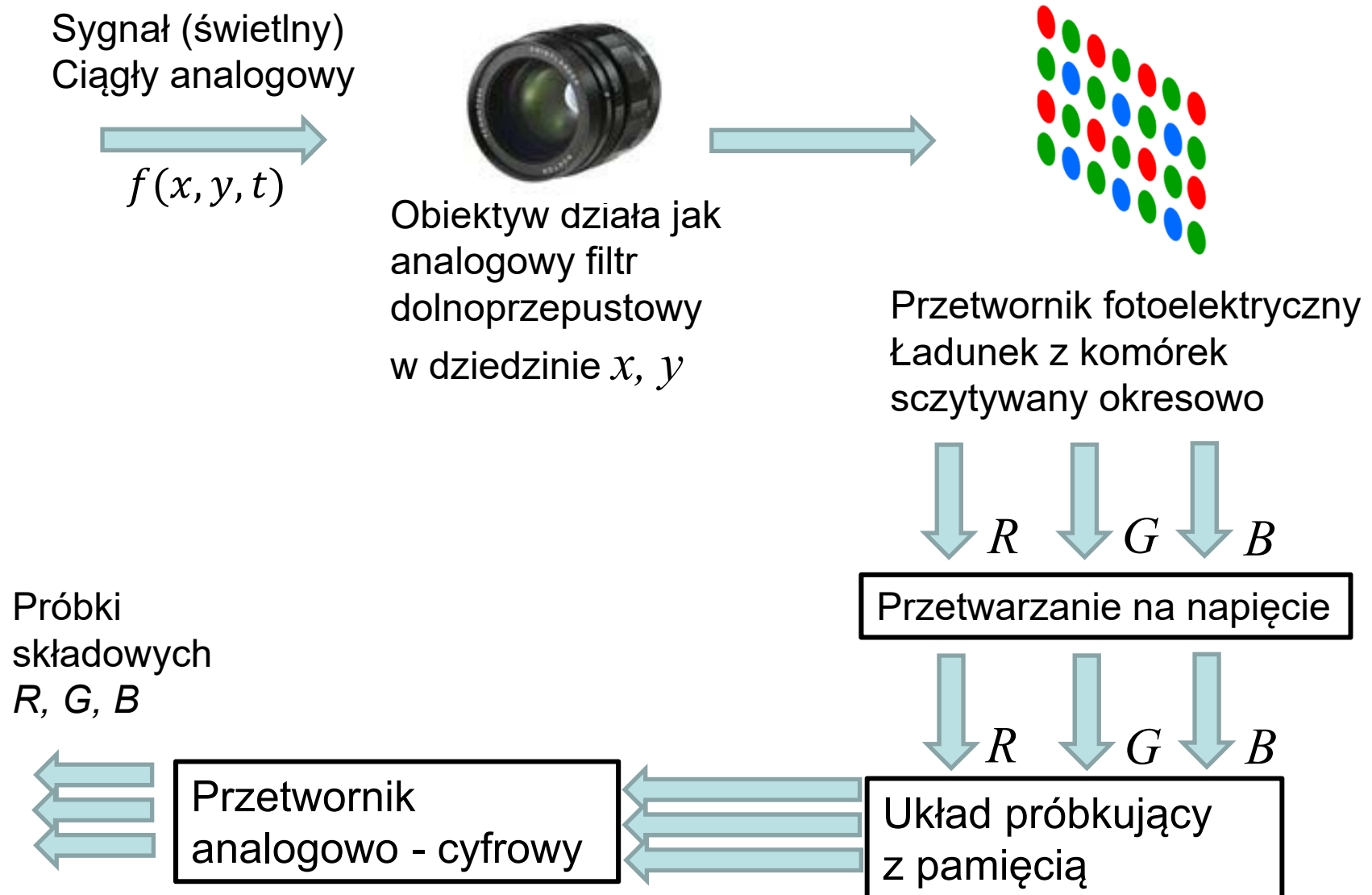


Charakterystyka częstotliwościowa =
Moduł transmitancji widmowej =
Moduł transformaty Fouriera
odpowiedzi impulsowej



Częstotliwość przestrzenna

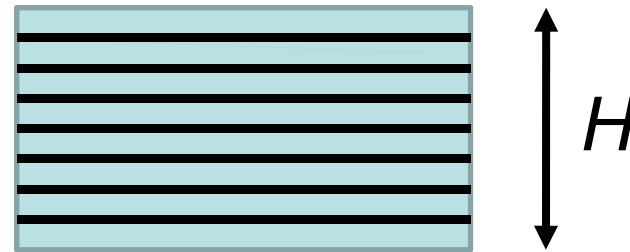
Akwizycja wizji w kamerze



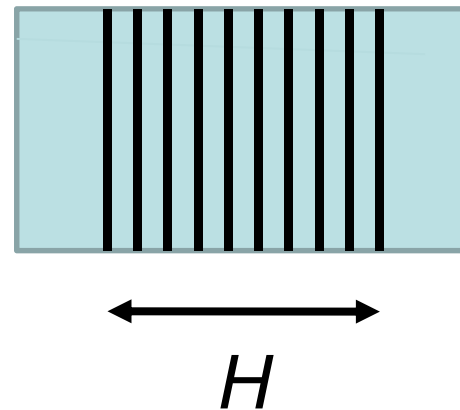
Kamery

Zdolność rozdzielcza albo rozdzielczość *Resolution*

Rozdzielczość
pionowa:



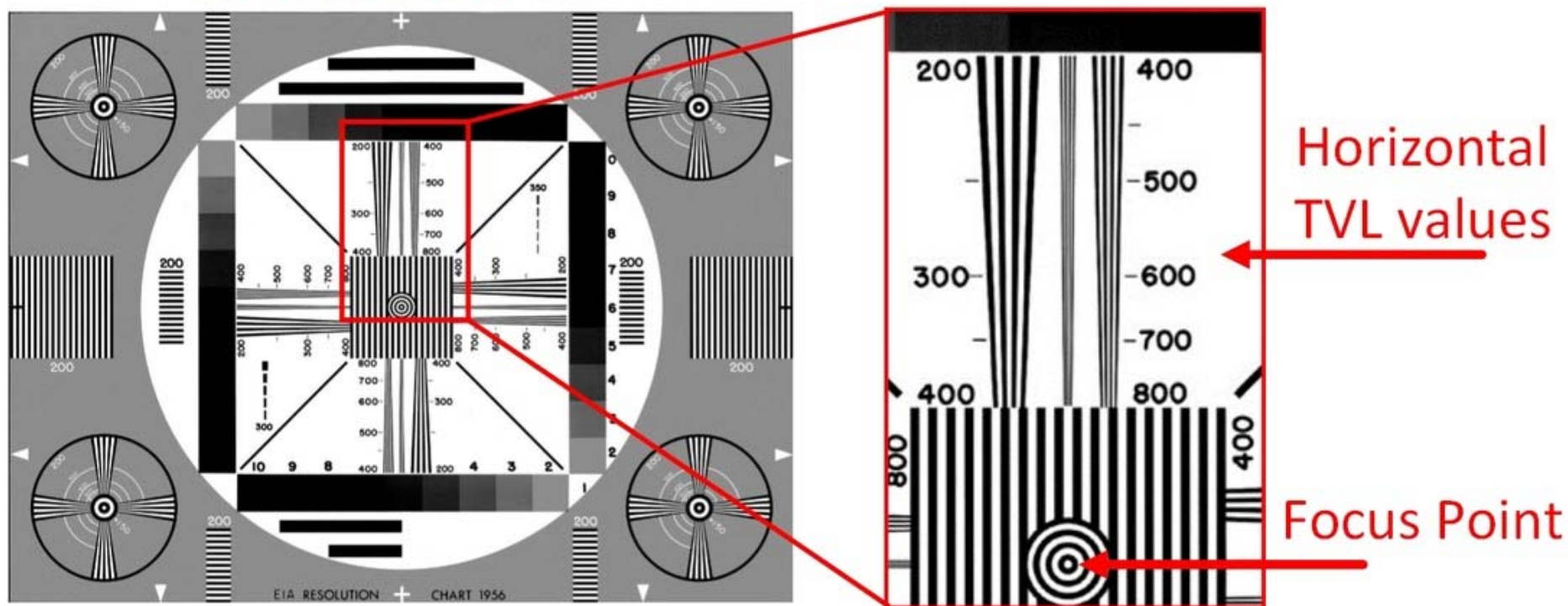
Rozdzielczość
pozioma:



Zdolność rozdzielcza albo rozdzielczość

Resolution

Aspect Ratio 4:3



2.2. Kamery

- Television
ENG – Electronic News Gathering
EFP – Electronic Field Production
- Camcoders



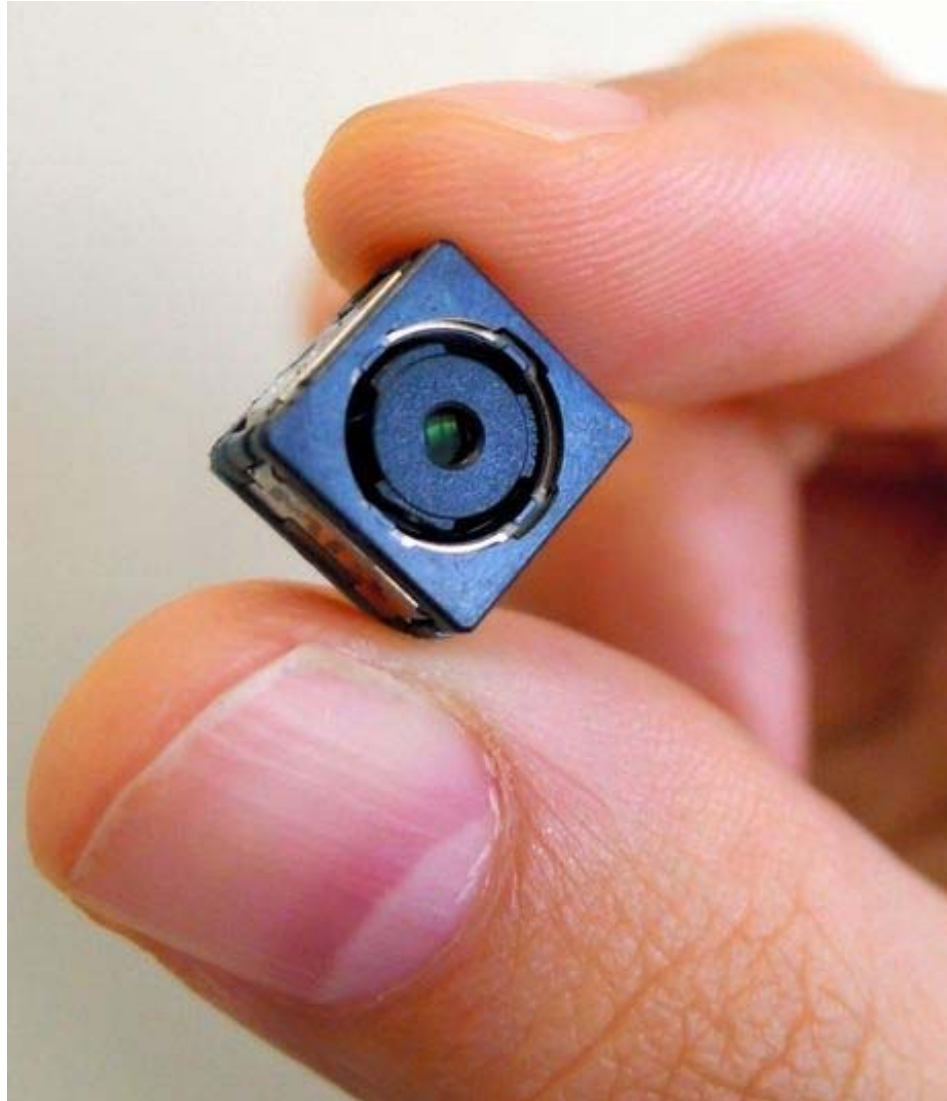
Movie camera



Studio cameras



Smartphone camera module



Computer vision, industrial applications



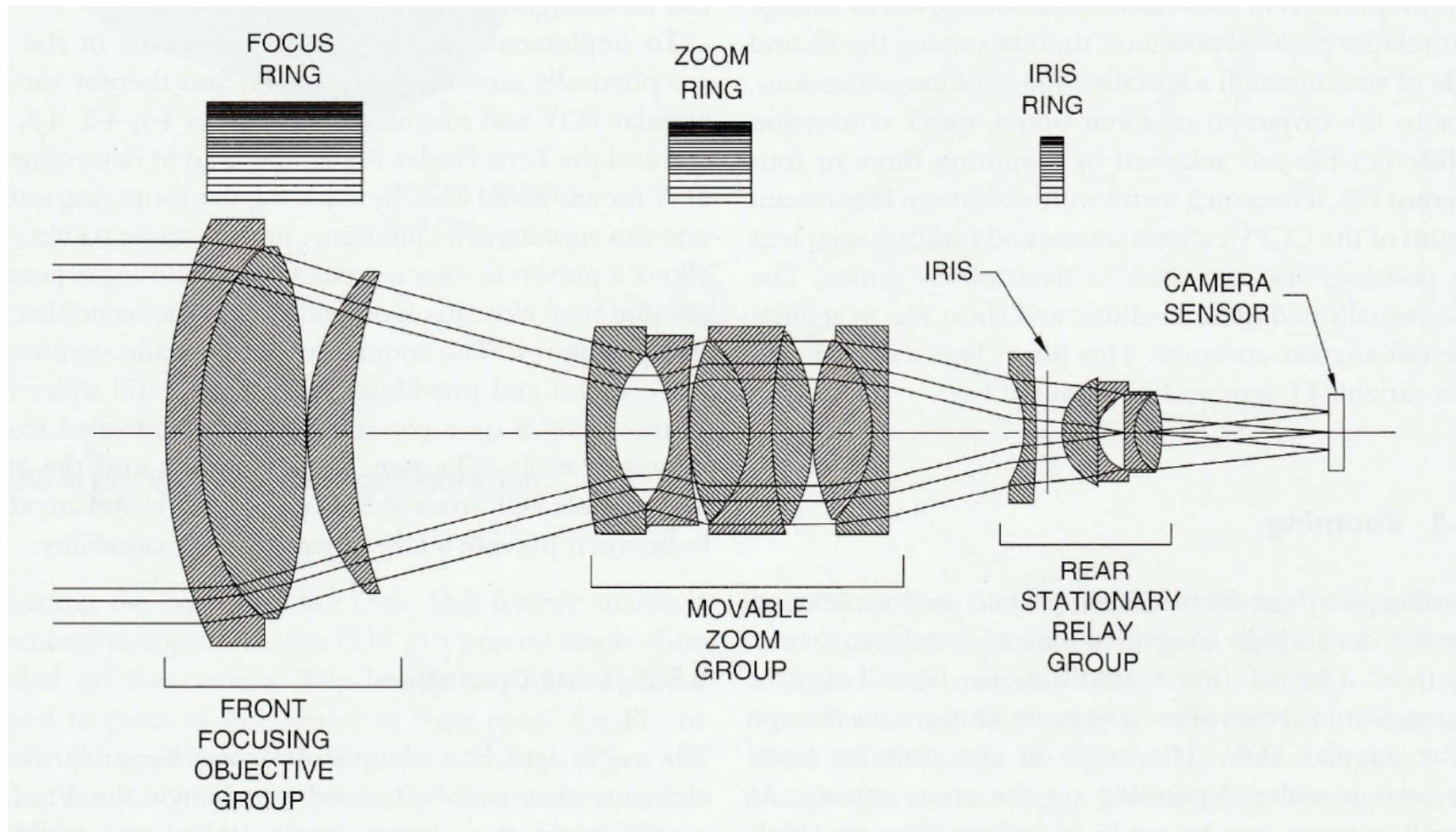
Kamery dozoru wizyjnego

Video surveillance cameras



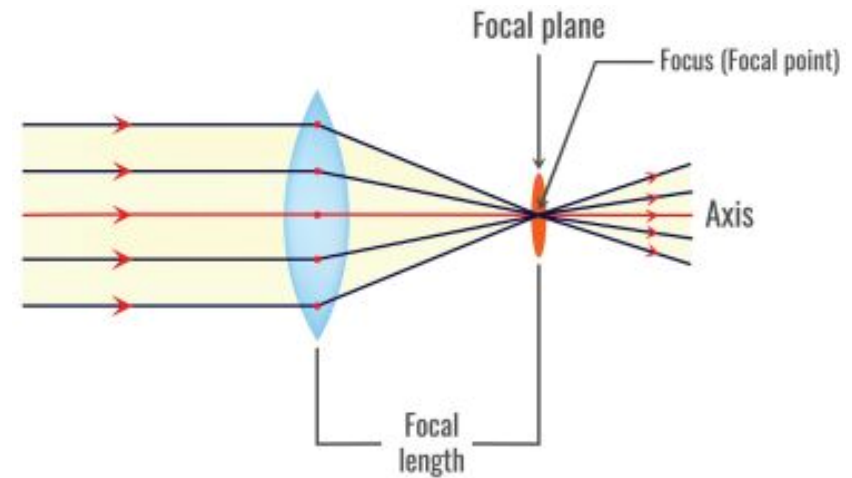
PTZ camera = Pan –Tilt – Zoom

Obiektyw - *lens*



Obiektyw

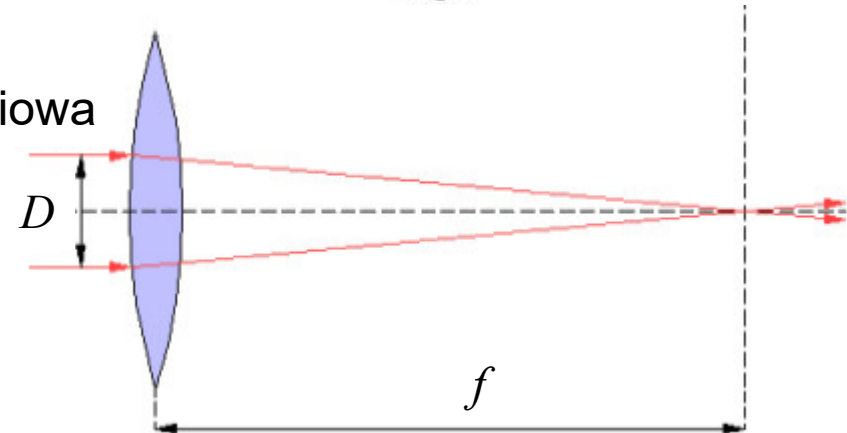
- Długość ogniskowej



- Przesłona, apertura

D – źrenica wejściowa

Zapis: $\frac{f}{f/D}$ np. $\frac{f}{1.5}$



- Charakterystyka w funkcji częstotliwości przestrzennych

2.3. Wizja i formaty obrazu

Proporcje obrazu

obraz elektroniczny	4 : 3	1,33 : 1
	16 : 9	1,78 : 1
		2,21 : 1
film	3 : 2	1,50 : 1
		1,66/1,85 : 1
		2,39 : 1

Format obrazu

Image format

1920 x 1080

liczba próbek w linii =
= liczba kolumn

number of samples in line

number of columns

liczba próbek w kolumnie =
= liczba linii

number of samples in column

number of lines

Przesyłanie wizji polami lub pełnymi obrazami

- Wizja analogowa polami (wybieranie międzyliniowe)
- Wizja cyfrowa SDTV polami
- Wizja cyfrowa HDTV 720linii pełnymi obrazami
- Wizja cyfrowa HDTV 1080linii polami
- Wizja cyfrowa UHD TV pełnymi obrazami
- IPTV, filmy w internecie pełnymi obrazami
- Kino cyfrowe pełnymi obrazami

Wybieranie międzyliniowe w sygnale cyfrowym Interlaced vs progressive video

Pole górne obrazu SD

Linia 0
Linia 2
Linia 4
⋮
Linia 574

Pole dolne obrazu SD

Linia 1
Linia 3
Linia 3
⋮
Linia 575

Linia 0 rejestruje obraz 20 ms wcześniej niż linia 1 !! (dla 25 obrazów na sekundę).

Linia 2 rejestruje obraz 20 ms wcześniej niż linia 3 !! (dla 25 obrazów na sekundę).

Itd.

Dla HD: 1080 linii pola w obrazie - 540 linii w każdym polu.

SDTV, HDTV, UHDTV

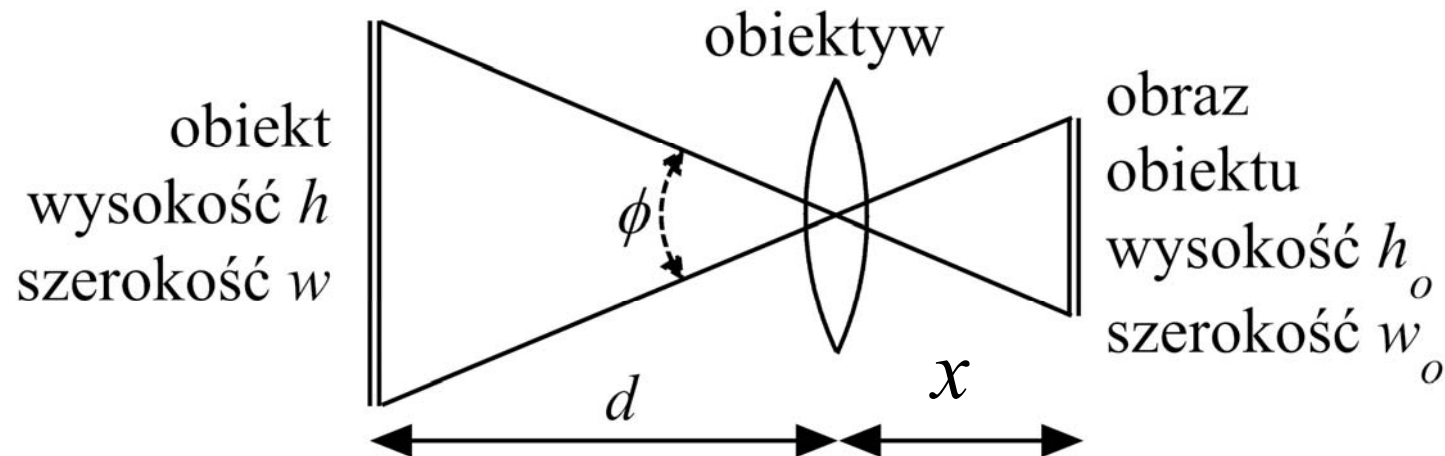
Formaty obrazów telewizyjnych w Europe	Częstotliwość obrazów	Liczba próbek w obrazie
SDTV Standard Definition Television 720 × 576	25 fps, w. międzyliniowe	< 750 000
HDTV High Definition TV 1280 × 720 1920 × 1080 (1440 × 1080)	25 or 50 fps, prog. (29.97/59.94) 25 fps, w. międzylin. (29.97)	> 750 000 < 3 000 000
UHDTV Ultra High Definition Television 3840 × 2160 7680 × 4320	50 / 100 fps, prog. (60/120) 50 / 100 fps, prog. (60/120)	> 3 000 000

Głębokość bitowa: SDTV, HDTV : 8 bit/sample
UHDTV : 10/12 bit/sample

Przykład

- Wyznaczyć prędkość bitową nieskompresowanej wizji HD (25 obraz/s) o liczbie linii 1080.

2.4. Tworzenie obrazu



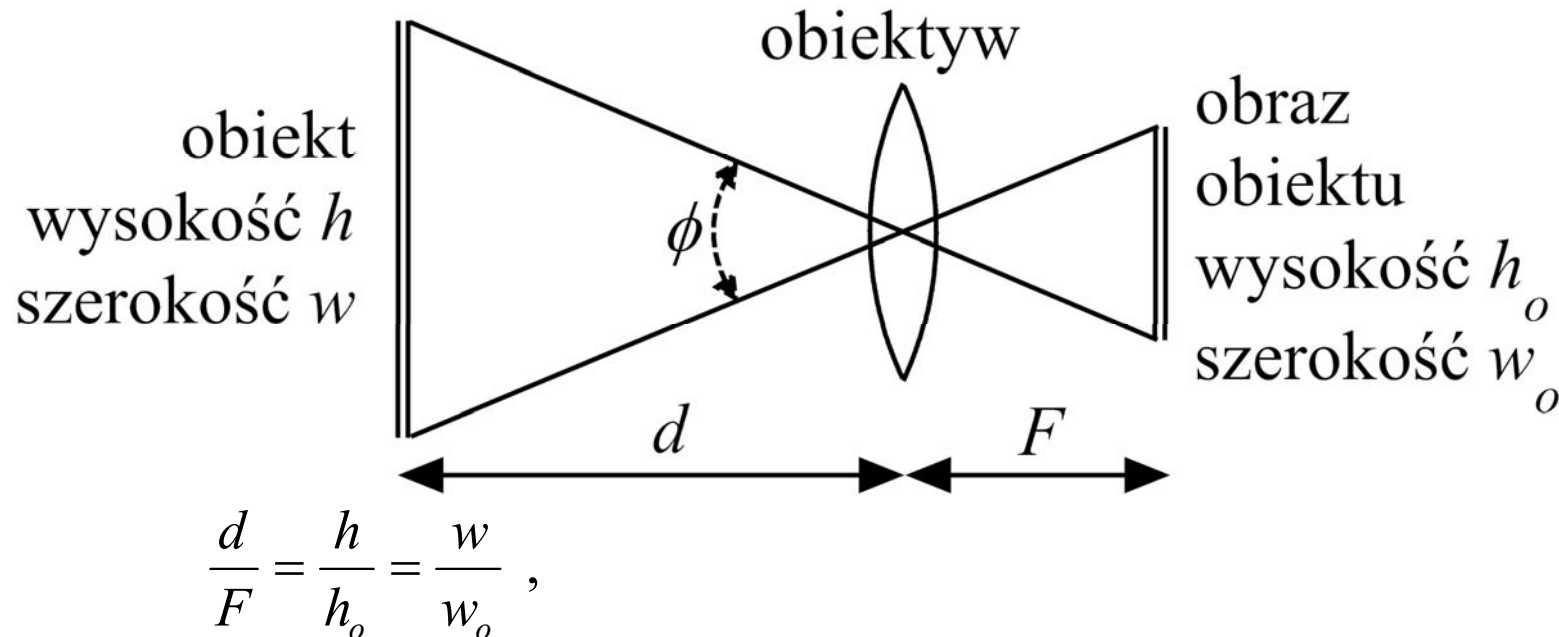
$$\frac{1}{d} + \frac{1}{x} = \frac{1}{F},$$

d – odległość między obiektem a obiektywem,
 x – odległość między obrazem a obiektywem,
 F – odległość ogniskowa.

Jeżeli d jest znacznie większe od $x \Rightarrow x \approx F$.

Tworzenie obrazu – model „kamery otworkowej”

Uproszczenie $x \approx F$



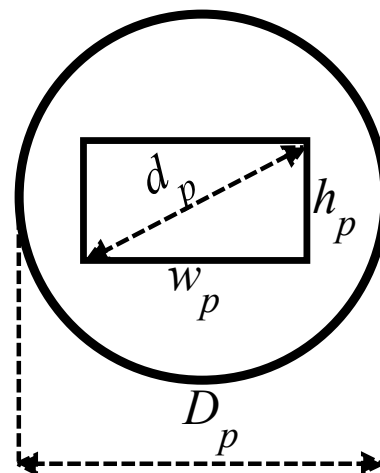
gdzie d – odległość między obiektem a obiektywem

F – odległość ogniskowa- *focal length*,

h, w – wysokość i szerokość obiektu ,

h_o, w_o – wysokość i szerokość obrazu obiektu.

Rozmiary przetworników optoelektronicznych



Typowe rozmiary półprzewodnikowych przetworników wizyjnych
Typical dimensions of semiconductor photoconverters

Nominalna średnica przetwornika w calach <i>Nominal converter diameter [inch]</i>	Proporcje rozmiarów obrazu: $w_p:h_p$ <i>Converter aspect ratio: $w_p:h_p$</i>	Rozmiary przetwornika <i>Converter size [mm]</i>		
		wysokość <i>high</i> h_p	szerokość <i>width</i> w_p	przekątna <i>diagonal</i> d_p
1	4:3	9,6	12,8	16,0
2/3	4:3	6,6	8,8	11,0
2/3	16:9	5,4	9,6	11,0
1/1,8	4:3	7,2	5,3	8,9
1/2	4:3	4,8	6,4	8,0
1/3	4:3	3,6	4,8	6,053
1/4	4:3	3,2	2,4	4,0

Przykład

- Monochromatyczna kamera dozoru wizyjnego posiada przetwornik o rozmiarze 1cm x 1 cm, na którym znajduje się 1000 x 1000 elementów światłoczułych.
- Kamera ma obserwować wejście do budynku o wysokości 2 metry znajdujące się w odległości 20 metrów od kamery.
- Dobrać obiektyw dla tej kamery wyznaczając jego długość ogniskowej tak, obraz wejścia stanowił 50% wysokości obrazu tworzonego w kamerze.

2.5. Barwy

2.5.1. Ludzki system widzenia *Human visual system (HVS)*

Promieniowanie świetlne widzialne

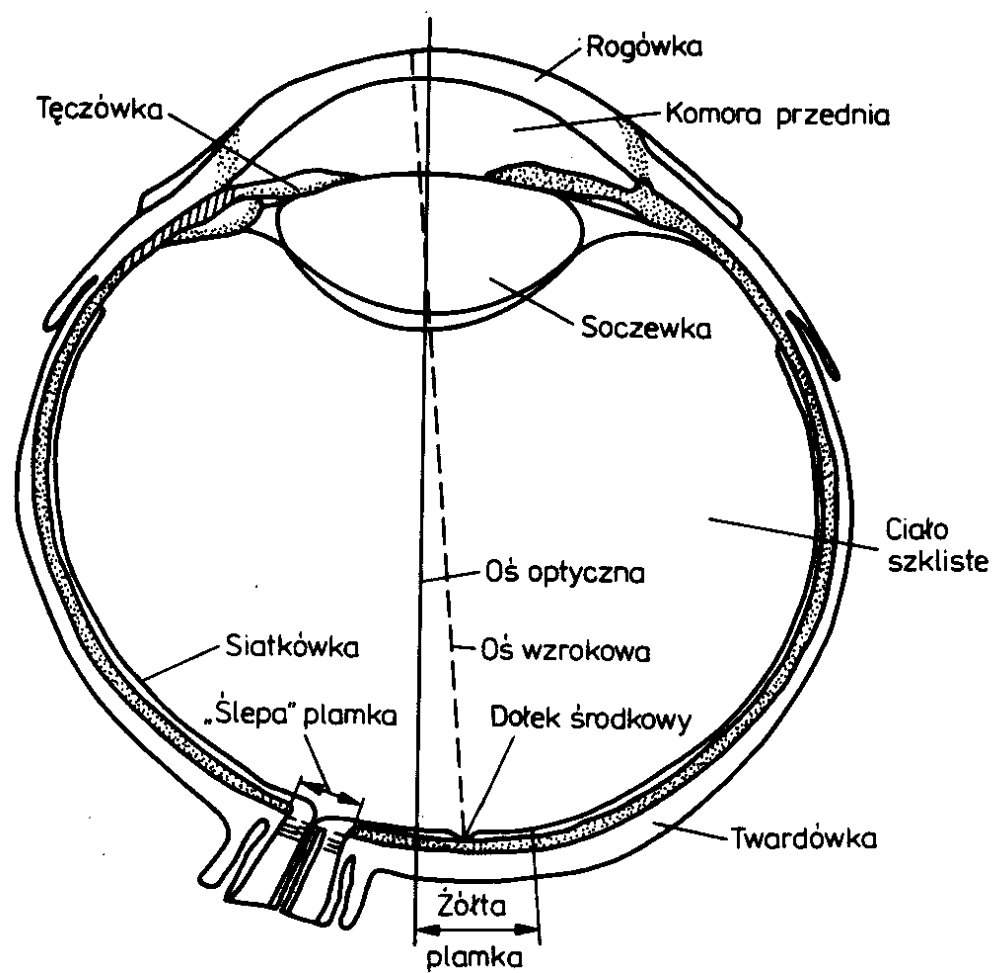
Fale elektromagnetyczne

$$f \approx 390 \div 830 \text{ THz} ,$$

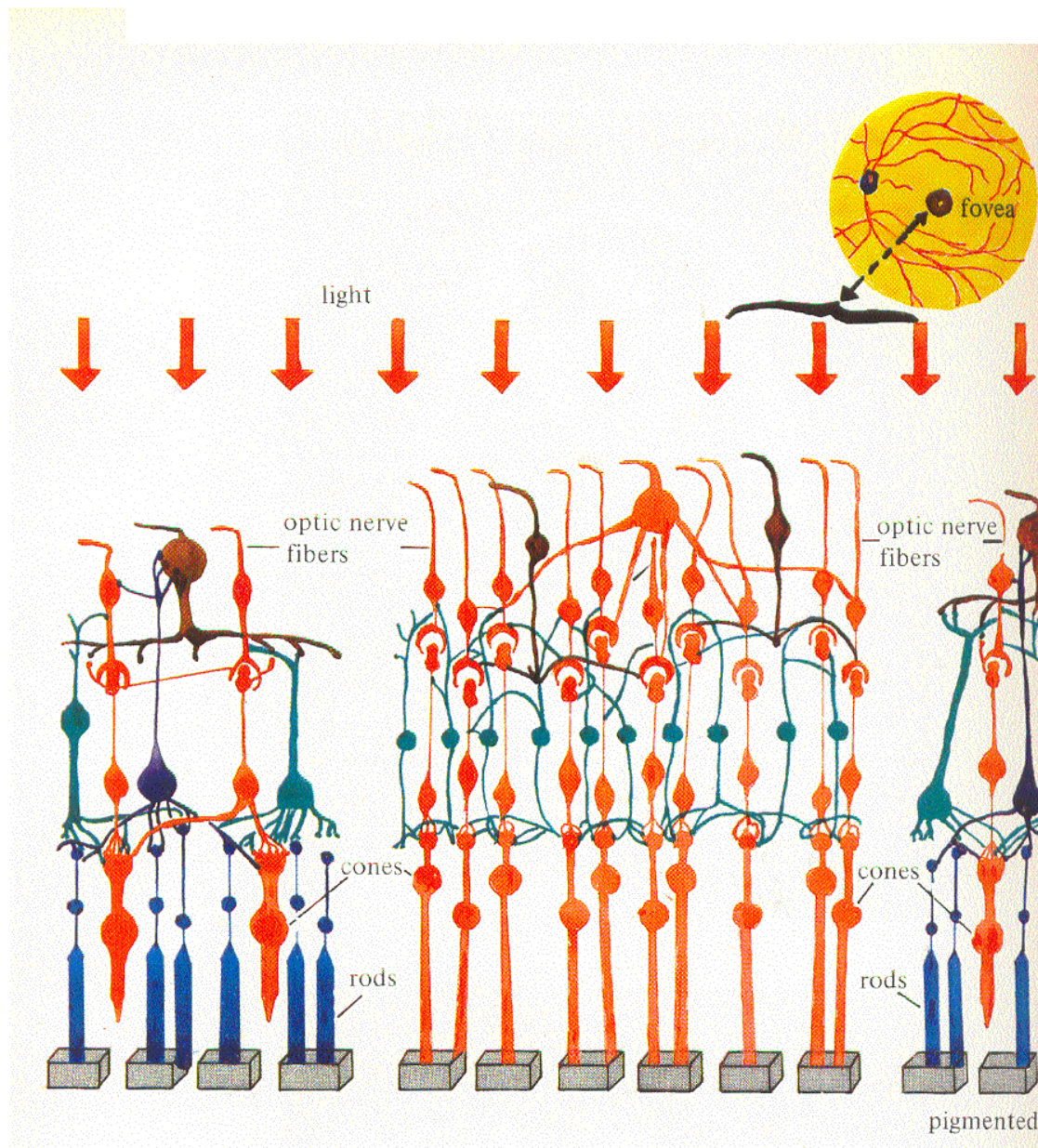
$$\lambda \approx \text{około } 380 \div 760 \text{ nm} .$$

Moc promieniowania - jednostka : wat [W]

Oko



Struktura siatkówki oka



Plamka żółta (*macula*)
średnica ok. 1,5 mm
i dołek środkowy (*fovea*)
średnica ok. 0,5 mm

Nerwy wzrokowe

Receptory:

Czopki (*cones*) – ok. 6 mln:
widzenie dzienne (fotopowe)

Pręciki (*rods*) – ok. 130 mln:
widzenie nocne (skotopowe)

FOTOMETRIA

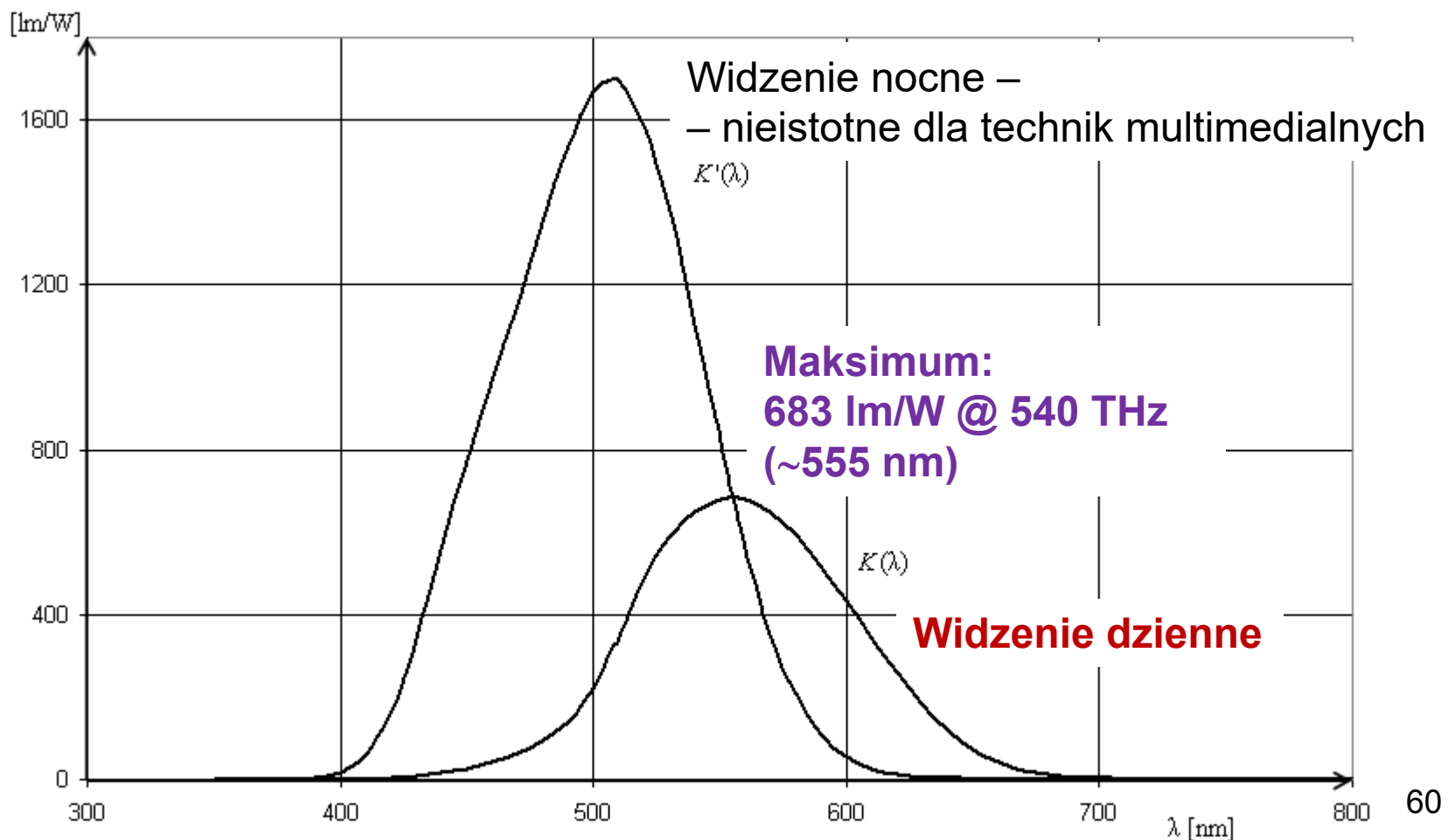
Strumień świetlny

- **Moc strumienia promieniowania P**
(*wielkość fizyczna*)
wyrażona w watach [W].
- Percepcja ilości światła
zależy od długości fali .
- *Przeliczenie z uwzględnieniem właściwości
systemu widzenia znormalizowanego
obserwatora.*
- **Strumień świetlny Φ**
(*wielkość psychofizyczna - fotometryczna*)
wyrażony w lumenach (oznaczenie: **lm**).

Skuteczność świetlna

Dla światła monochromatycznego: $\Phi = K(\lambda) \cdot P$

$K(\lambda)$ - skuteczność świetlna [lm / W]



Skuteczność świetlna

Dla światła będącego mieszaniną światel o różnych długościach fali świetlnej:

$$\Phi = \int_{380 \text{ nm}}^{760 \text{ nm}} (K(\lambda) \cdot p(\lambda)) d\lambda$$

$K(\lambda)$ - skuteczność świetlna [lm / W]

$p(\lambda)$ - gęstość widmowa mocy promieniowania

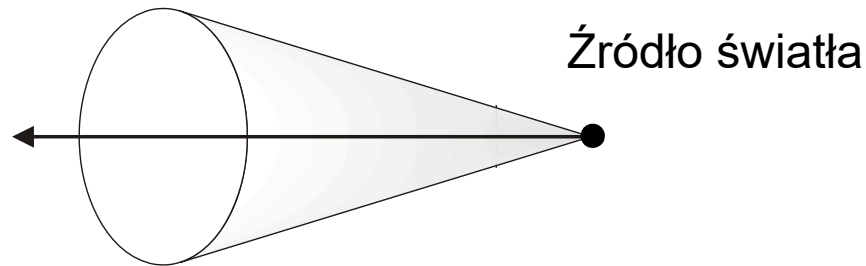
$$P = \int_{380 \text{ nm}}^{760 \text{ nm}} p(\lambda) d\lambda \text{ - całkowita moc promieniowania}$$

Światłość

to

strumień świetlny promieniowany przez źródło światła w jednostkowy kąt bryłowy.

$$I = \frac{d\Phi}{d\Omega}$$



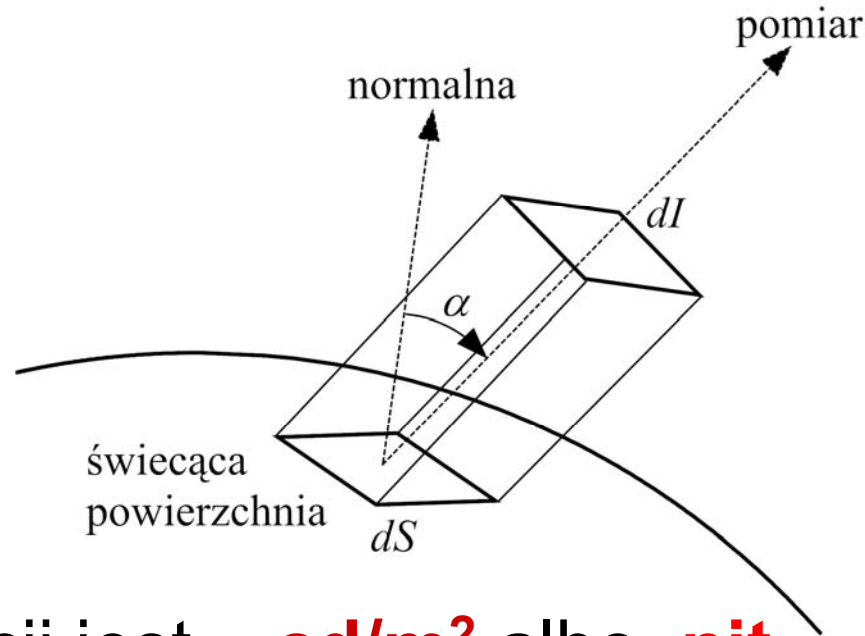
Jednostkowy
kąt bryłowy

= Jest to kąt bryłowy o wierzchołku w środku kuli, wycinający z powierzchni tej kuli pole równe kwadratowi jej promienia.

Jednostką światłości jest **kandela** [cd]
(jednostka podstawowa układu SI).

Luminancja

$$L = \frac{dI}{dS \cdot \cos \alpha}$$



- Jednostką luminancji jest **cd/m²** albo **nit**.
- Np.: Maksymalna wartość luminancji charakteryzuje monitor :
np. dla SDR ok. 400 cd/m² to spora wartość
(SDR – Standard Dynamic Range).

Widzenie barwne

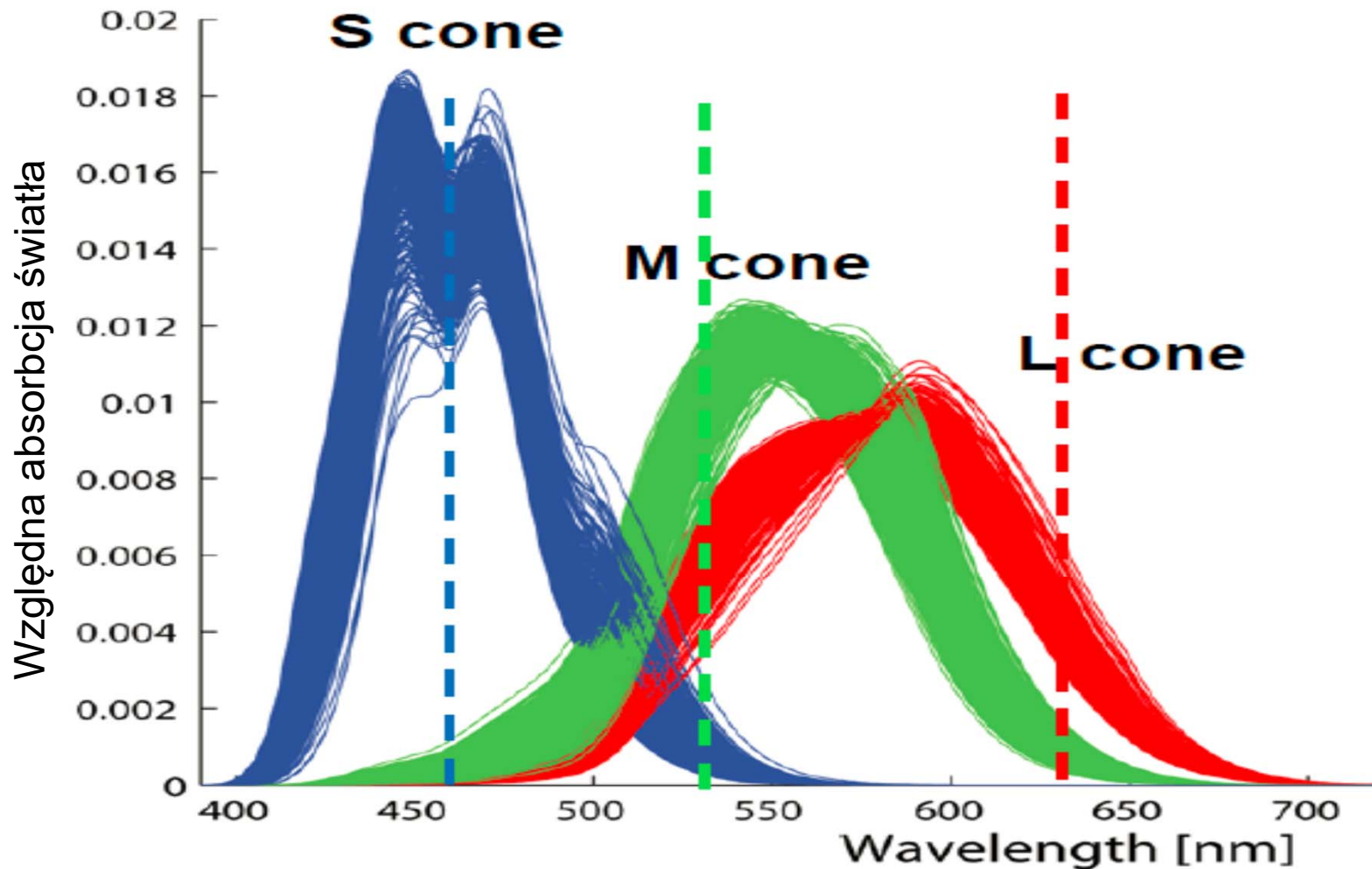
Postrzegana barwa światła zależy od jego rozkładu widmowego, czyli od proporcji udziału promieniowania o różnych długościach fali.

Różne barwy odpowiadają różnym rozkładom widmowym.

3 typy czopków

Czopki typu S, M, L

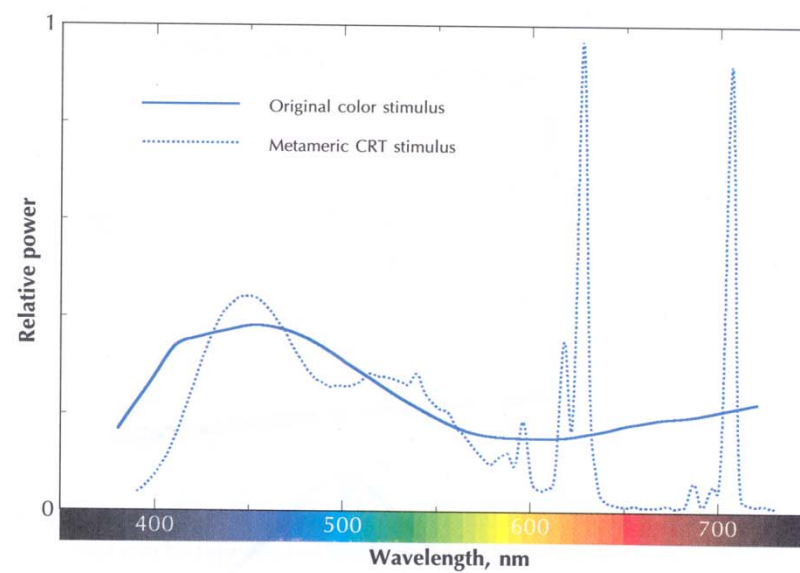
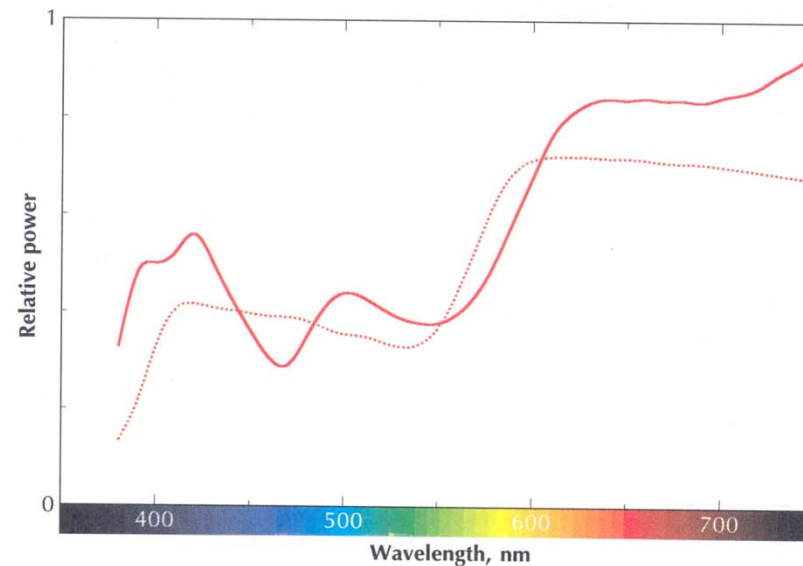
Zebrane krzywe dla 1000 osób



Długość fali światła)

Właściwość wzroku:

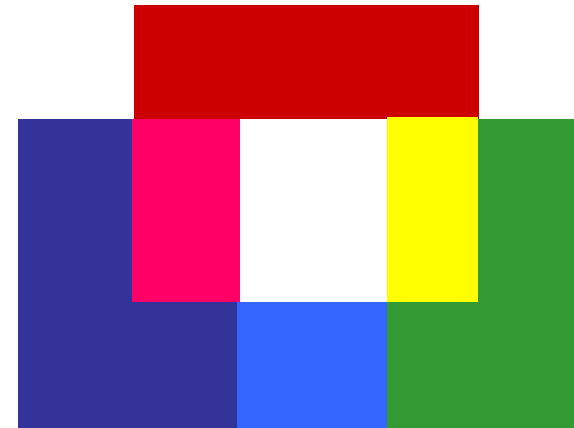
Jednej postrzeganej barwie odpowiada wiele różnych rozkładów widmowych.



2.5.2. Reprezentacja barw za pomocą trzech barw podstawowych

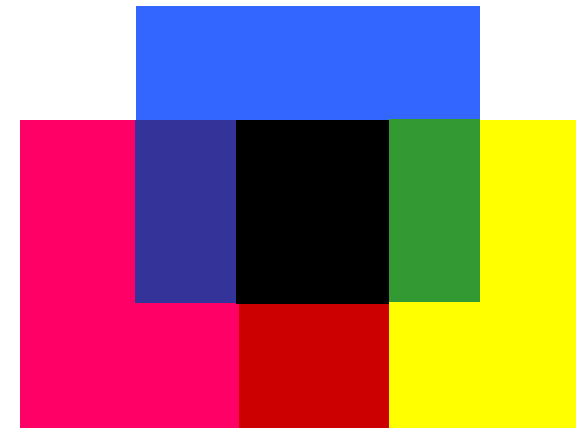
Mieszanie addytywne

czzerwona (red – R),
zielona (green – G),
niebieska (blue – B) .



Mieszanie subtraktywne

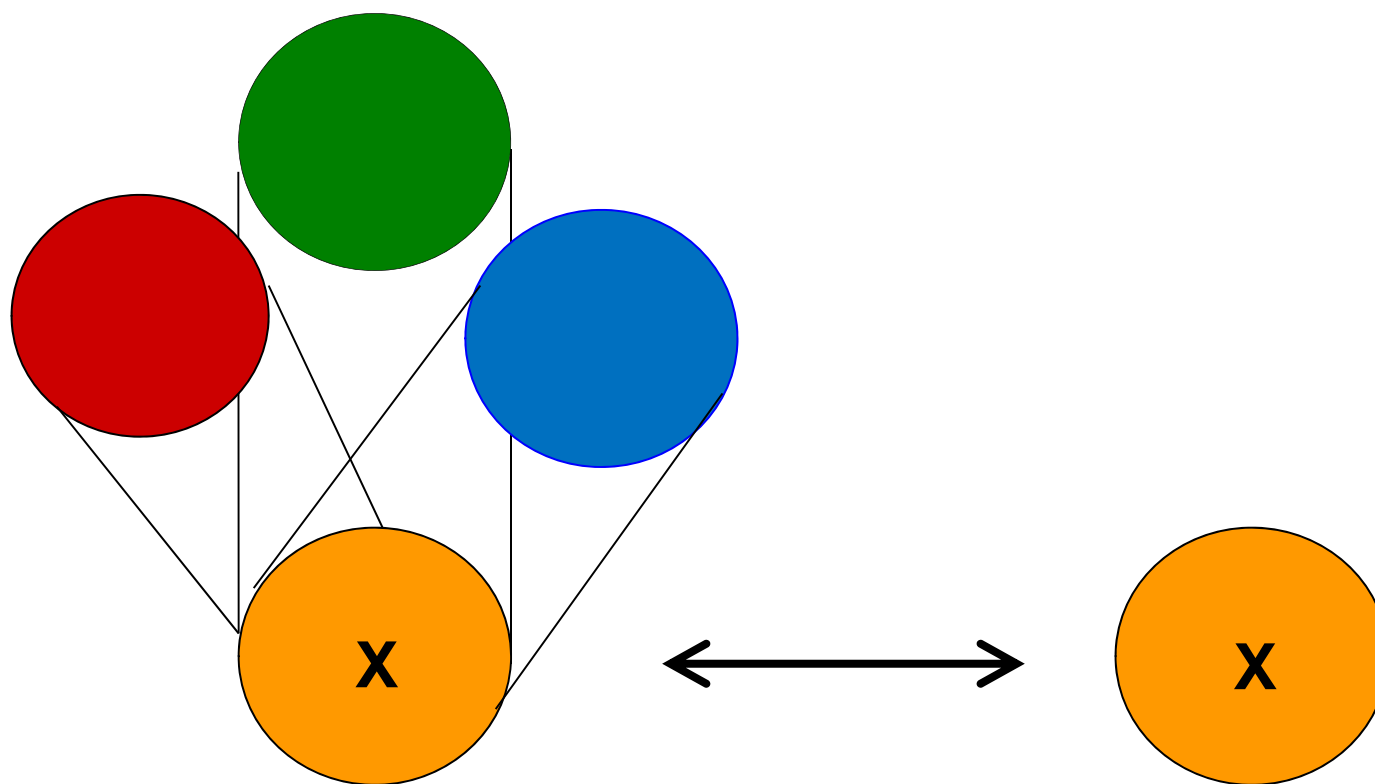
błękitna (C- cyan)
purpurowa (M - magenta)
żółta (Y - yellow)



Światła barwne możemy przedstawiać jako sumy światel o barwach podstawowych R, G, B

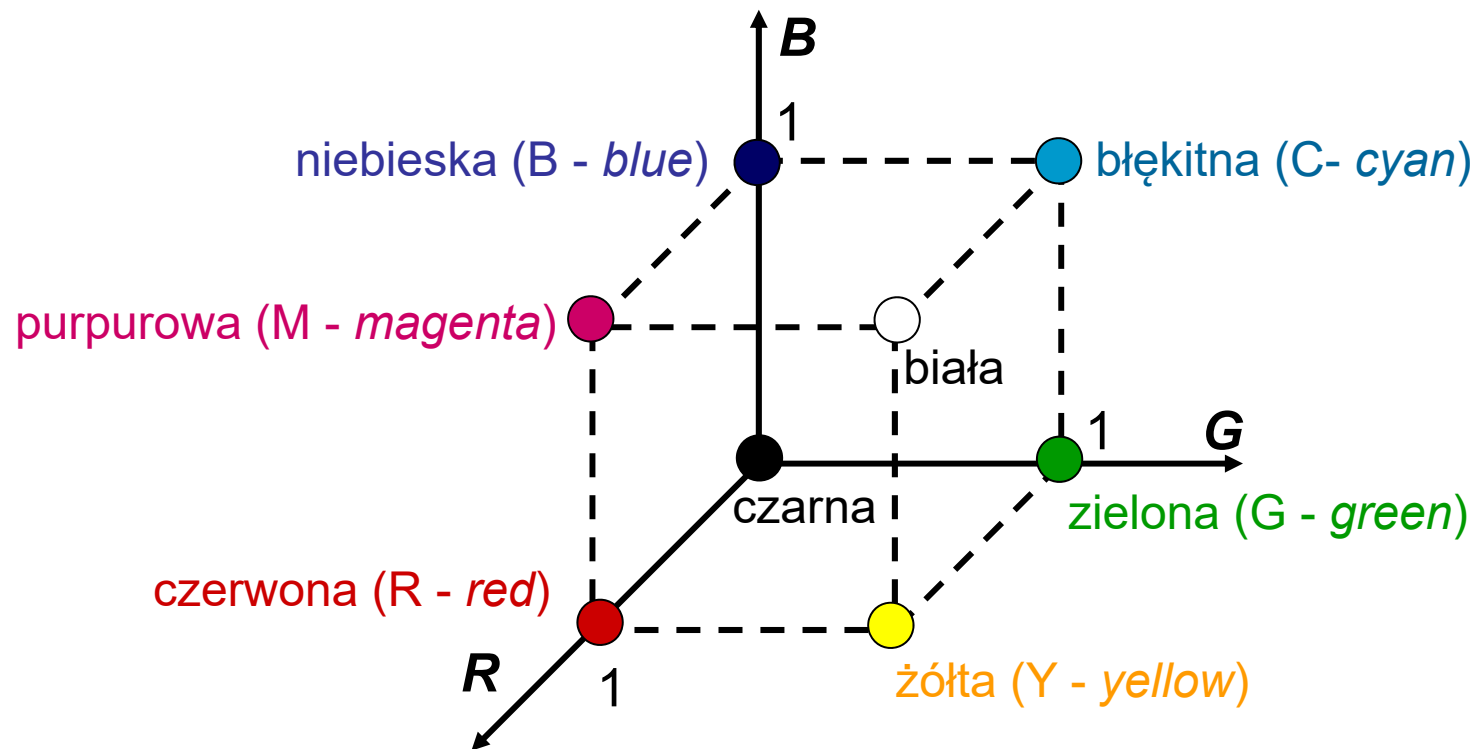
Strumienie świetlne:

$$X = R + G + B \quad [\text{lm}]$$



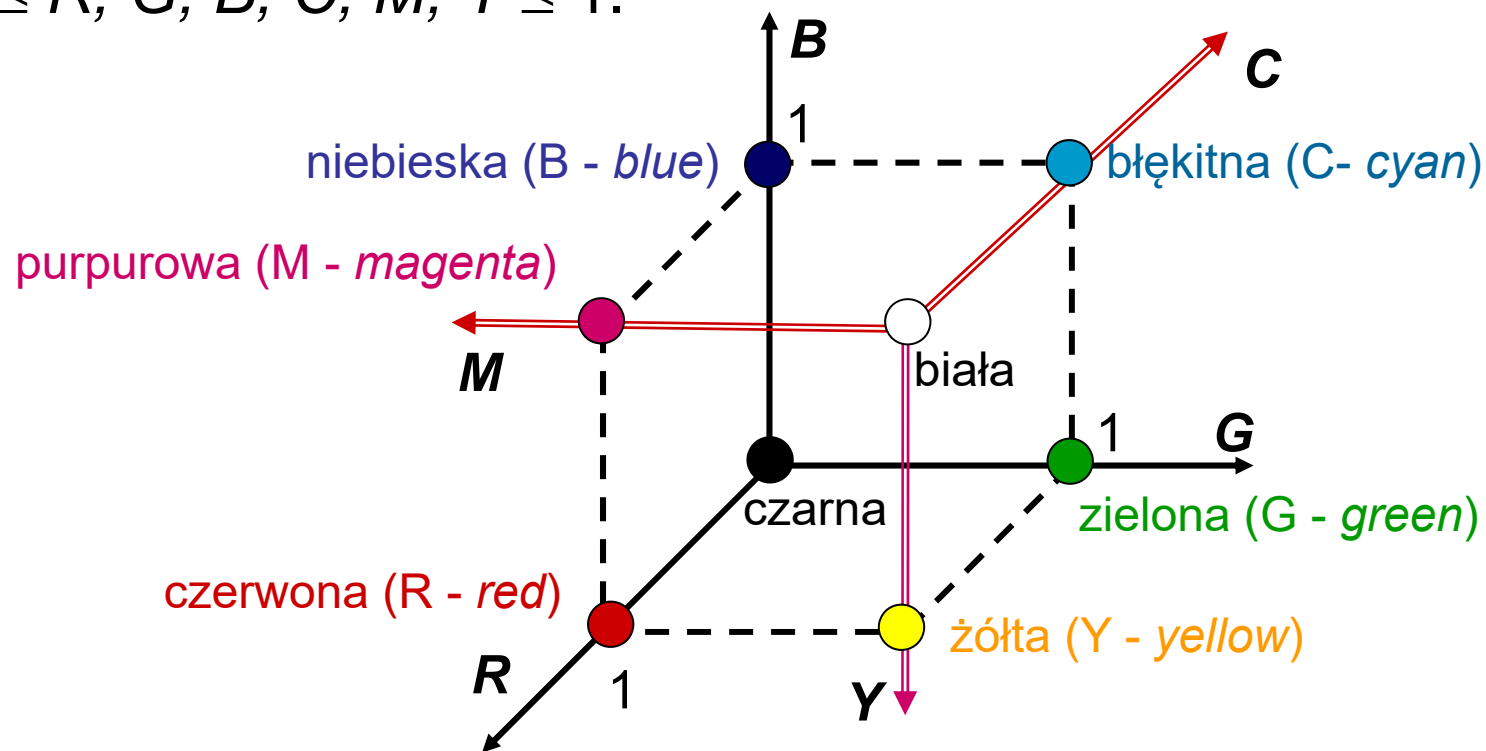
Światła barwne można przedstawiać jako trójki (R, G, B) lub (C, M, Y)

$$0 \leq R, G, B, C, M, Y \leq 1.$$



Zależności pomiędzy reprezentacjami (*R,G,B*) oraz (*C,M,Y*)

$$0 \leq R, G, B, C, M, Y \leq 1.$$



$$\begin{aligned}C &= 1 - R, \\M &= 1 - G, \\Y &= 1 - B.\end{aligned}$$

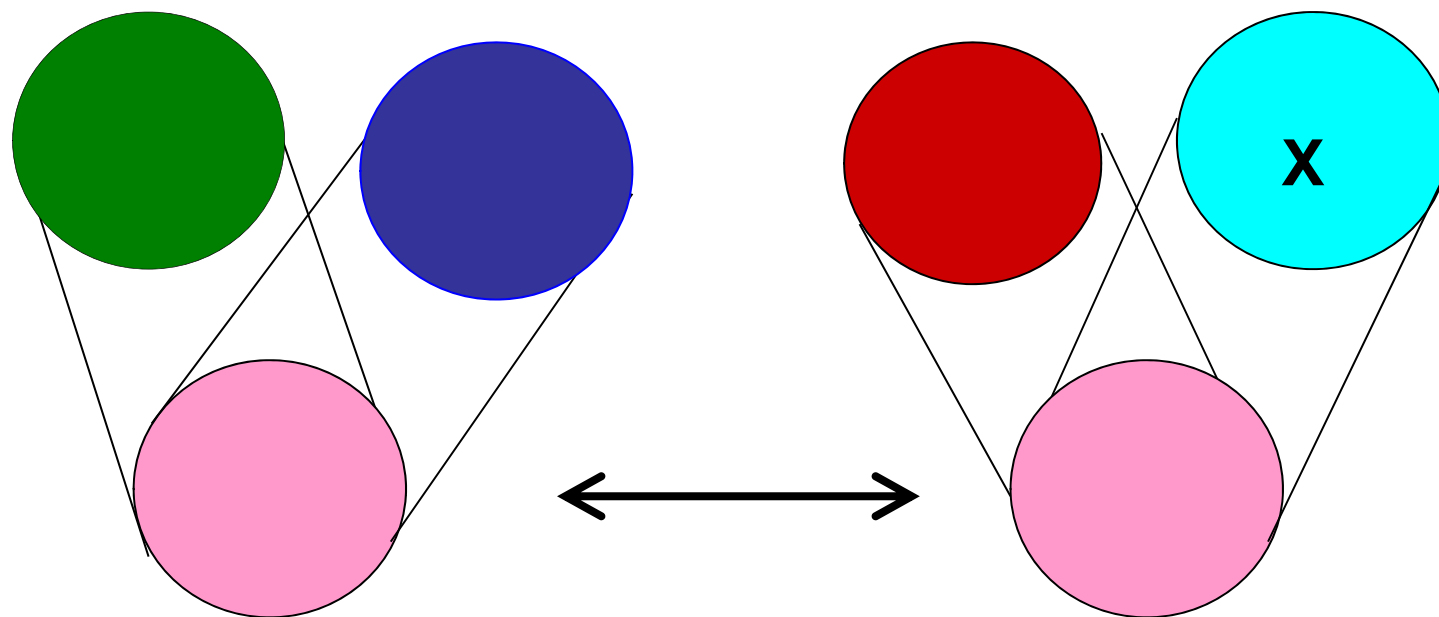
$$\begin{aligned}R &= 1 - C, \\G &= 1 - M, \\B &= 1 - Y.\end{aligned}$$

Ułomność prostego modelu RGB

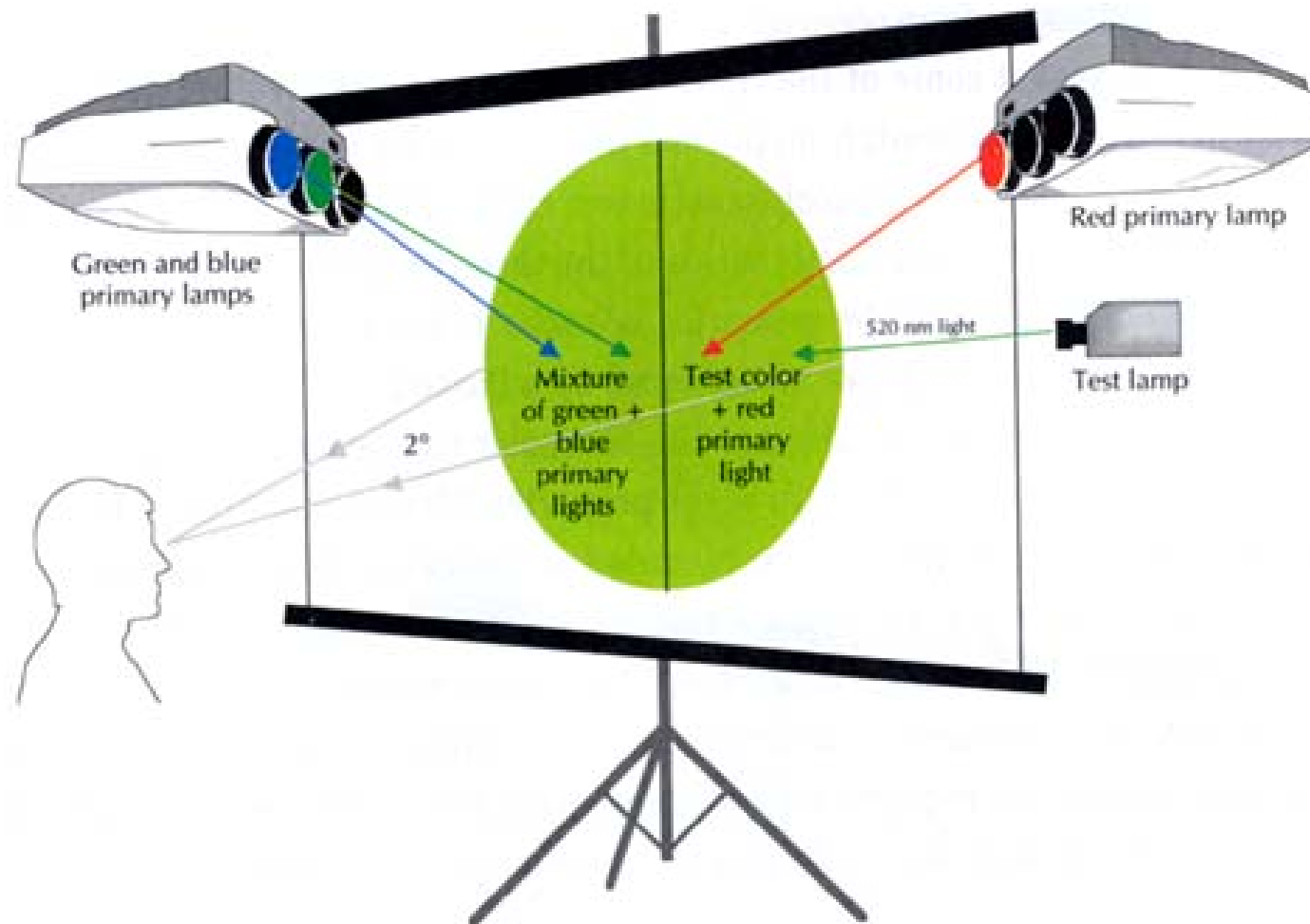
Ujemny udział składowej R

$$G + B = R + X$$

$$G + B - R = X$$



Eksperymentalne wyznaczanie rozkładów tróchromatycznych



Wniosek

- Istnieją barwy widzialne, których nie można odtworzyć przez addytywne mieszanie barw podstawowych R, G, B.

2.5.3. Układ XYZ

Przestrzeń barw, w której addytywne mieszanie barw podstawowych pozwala na uzyskanie wszystkich barw widzialnych, a udziały wszystkich barw podstawowych są zawsze nieujemne.

X, Y, Z – trzy urojone barwy podstawowe – nie odpowiadają barwom widzialnym.

Przestrzeń XYZ jest abstrakcją matematyczną.

Przestrzeń XYZ jest podstawową przestrzenią barw w kolorymetrii.

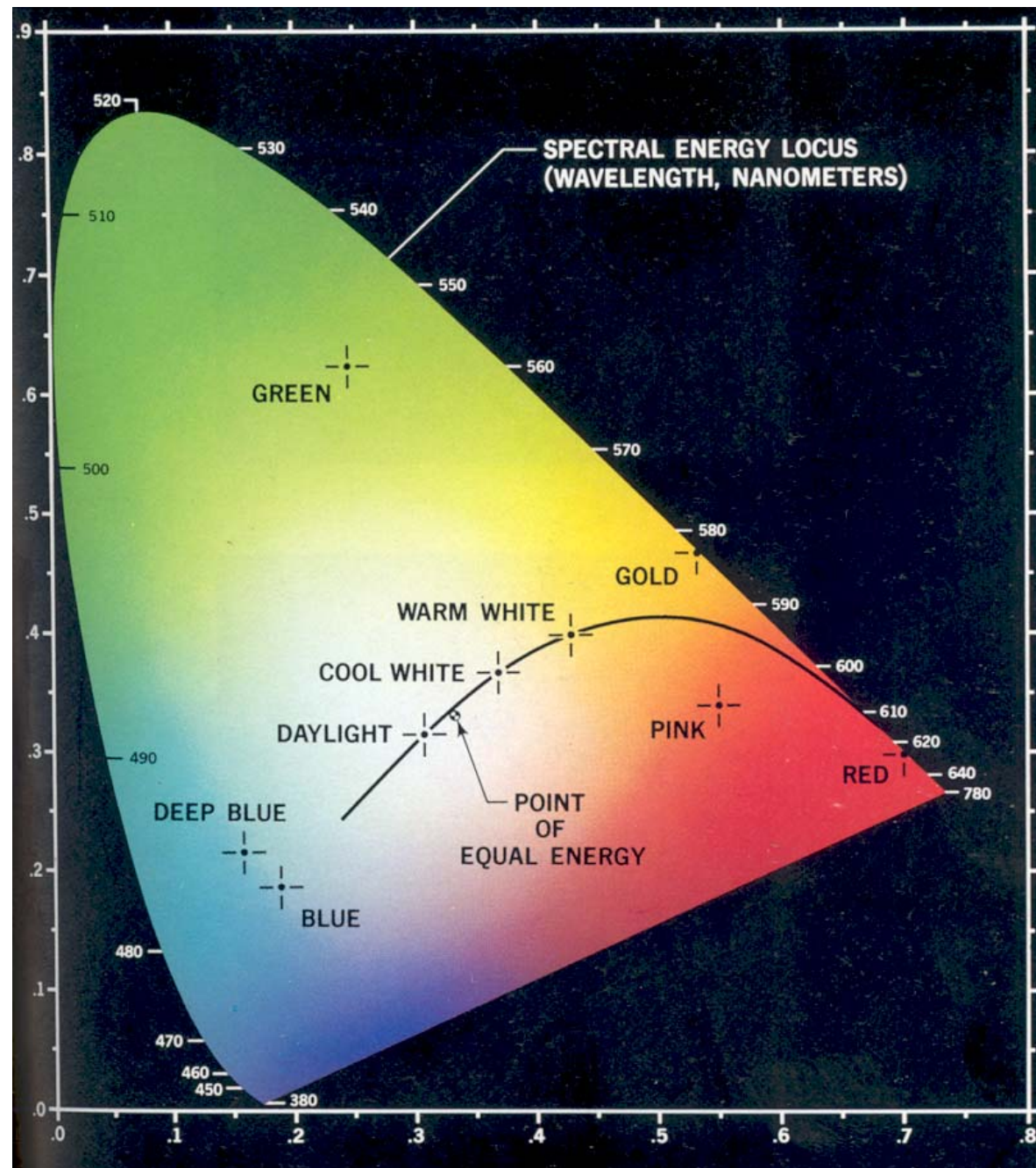
Współrzędne trójkromatyczne *xyz*

$$x = X/(X+Y+Z),$$

$$y = Y/(X+Y+Z),$$

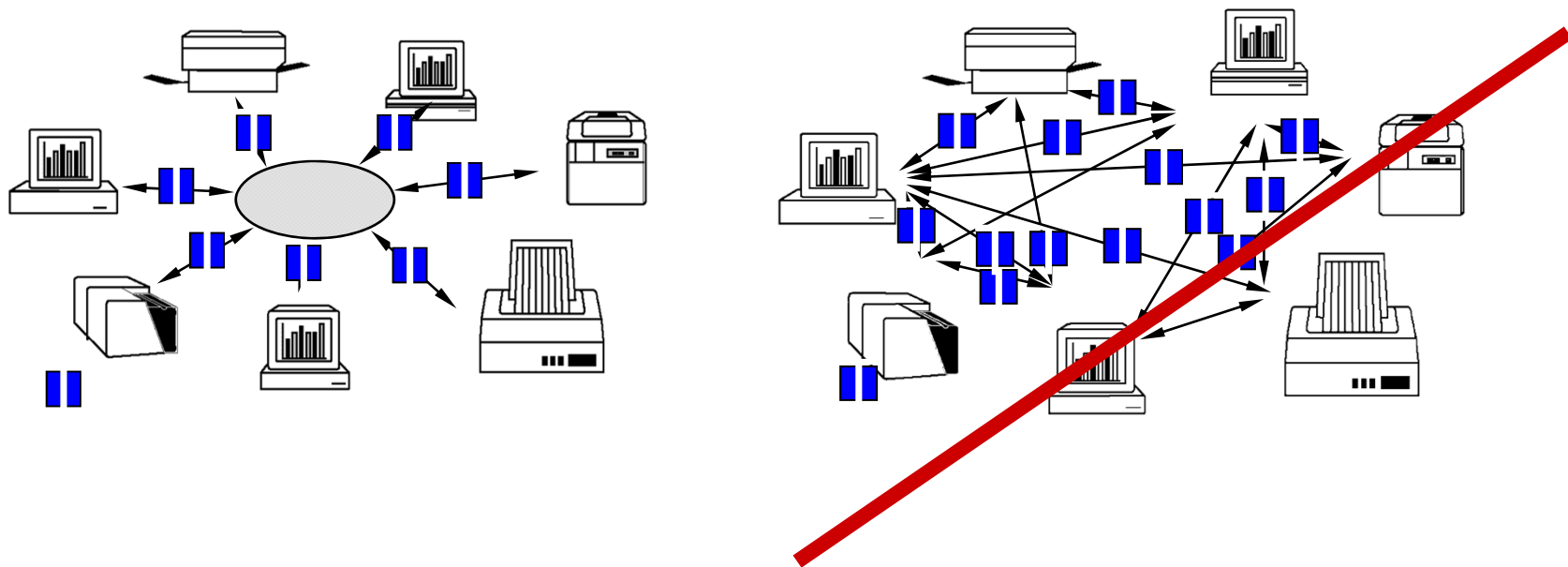
$$z = Z/(X+Y+Z).$$

$$0 \leq x, y, z \leq 1, \quad x + y + z = 1$$



Znaczenie układu XYZ

- Układ XYZ jest uniwersalnym układem odniesienia dla różnych reprezentacji barw . Używana także w dokumentacji urządzeń (monitorów, drukarek, itd.).
- Używany jako tzw. „profil barwny” pośredni $\Rightarrow$ redukcja liczby sterowników urządzeń.



2.5.4. Odmiany układu RGB

Przykładowe warianty układu RGB używane w technice obecnie lub w przeszłości

	R		G		B	
	x	y	x	y	x	y
NTSC (FCC) (Telewizja analogowa w USA)	0,670	0,330	0,210	0,710	0,140	0,080
EBU (Telewizja analogowa w Europie)	0,640	0,330	0,290	0,600	0,150	0,060
Adobe <i>RGB</i>	0,640	0,330	0,210	0,710	0,150	0,060
SMPTE RP 145	0,630	0,340	0,310	0,595	0,155	0,070
BT.601 (SDTV) Wersja 625 (Europa)	0,640	0,330	0,290	0,600	0,150	0,060
BT.601 (SDTV) Wersja 525 (USA)	0,630	0,340	0,310	0,595	0,155	0,070
BT.709/ <i>sRGB</i> (HDTV)	0,640	0,330	0,300	0,600	0,150	0,060
BT.2020 (UHDTV)	0,708	0,292	0,170	0,797	0,131	0,046

Przykład : Konwersja pomiędzy RGB-RT.709 i RGB EBU

$$R_e = 0.9578 * R_7 + 0.0422 * G_7 + 0.0000 * B_7$$

$$G_e = 0.0000 * R_7 + 1.0000 * G_7 + 0.0000 * B_7$$

$$B_e = 0.0000 * R_7 + 0.0118 * G_7 + 0.9882 * B_7$$

$$R_7 = 1.0440 * R_e - 0.0440 * G_e + 0.0000 * B_e$$

$$G_7 = 0.0000 * R_e + 1.0000 * G_e + 0.0000 * B_e$$

$$B_7 = 0.0000 * R_e - 0.0119 * G_e + 1.0119 * B_e$$

Przykład : Konwersja z RGB BT.709 do XYZ

$$X = 0.412 * R + 0.358 * G + 0.180 * B$$

$$Y = 0.213 * R + 0.715 * G + 0.072 * B$$

$$Z = 0.019 * R + 0.119 * G + 0.950 * B$$

$$R = 3.241 * X - 1.537 * Y - 0.499 * Z$$

$$G = -0.969 * X + 1.876 * Y + 0.042 * Z$$

$$B = 0.056 * X - 0.204 * Y + 1.057 * Z$$

Liniowe transformacje barw

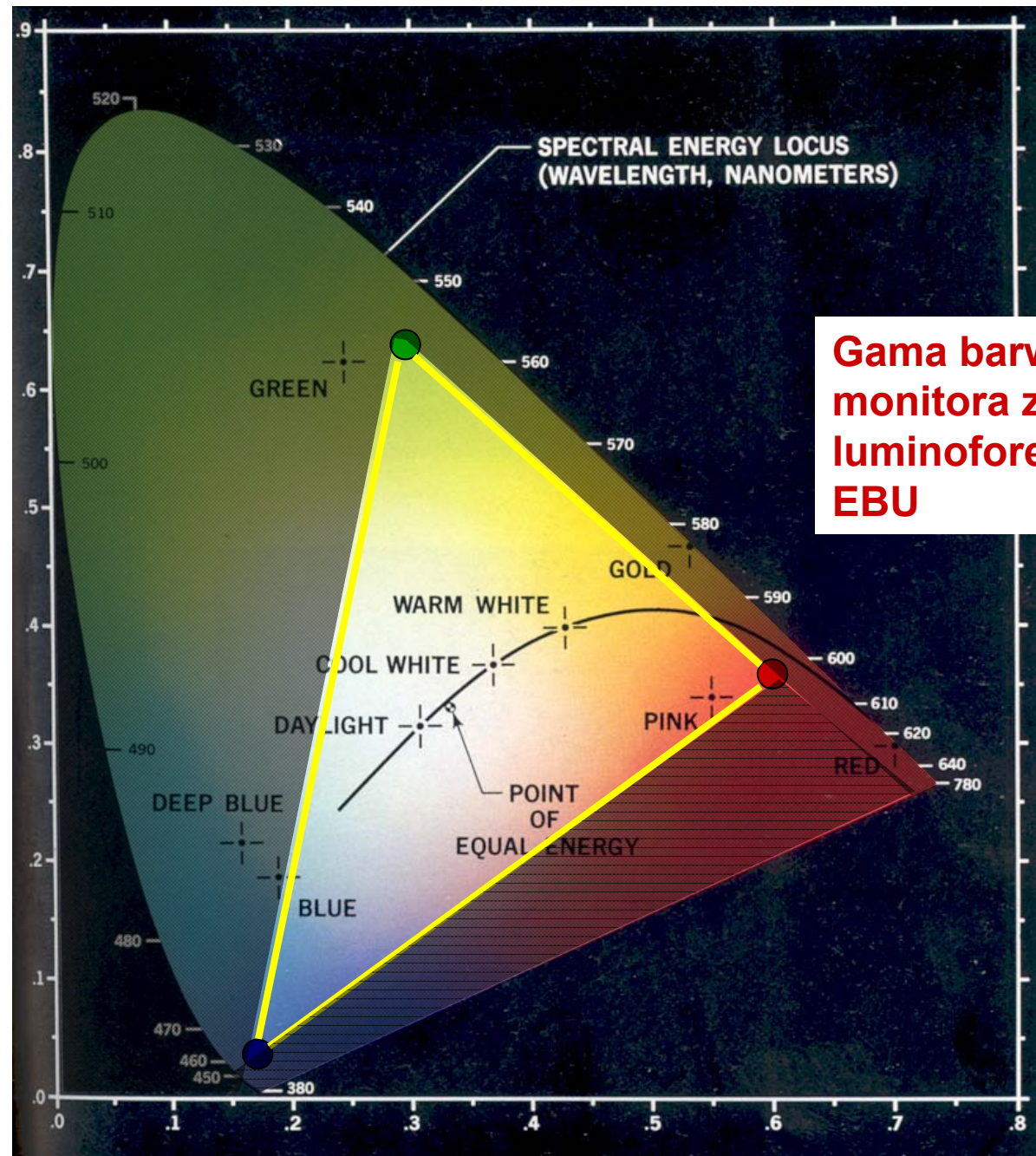
Ogólnie

$$\mathbf{B} = \mathbf{T} \cdot \mathbf{A}$$

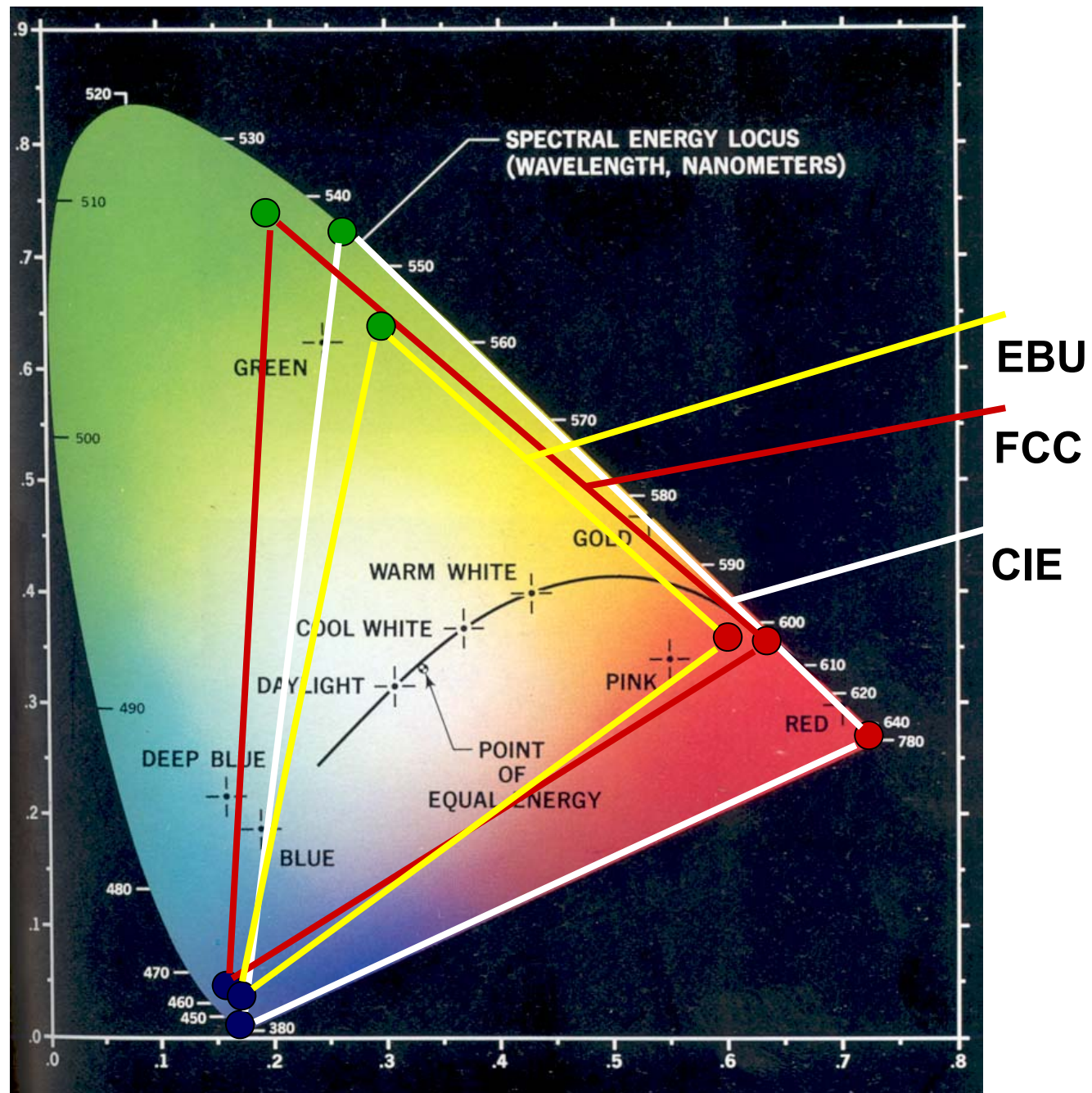
$$\mathbf{A} = \mathbf{T}^{-1} \cdot \mathbf{B}$$

2.5.5. Gamy barw

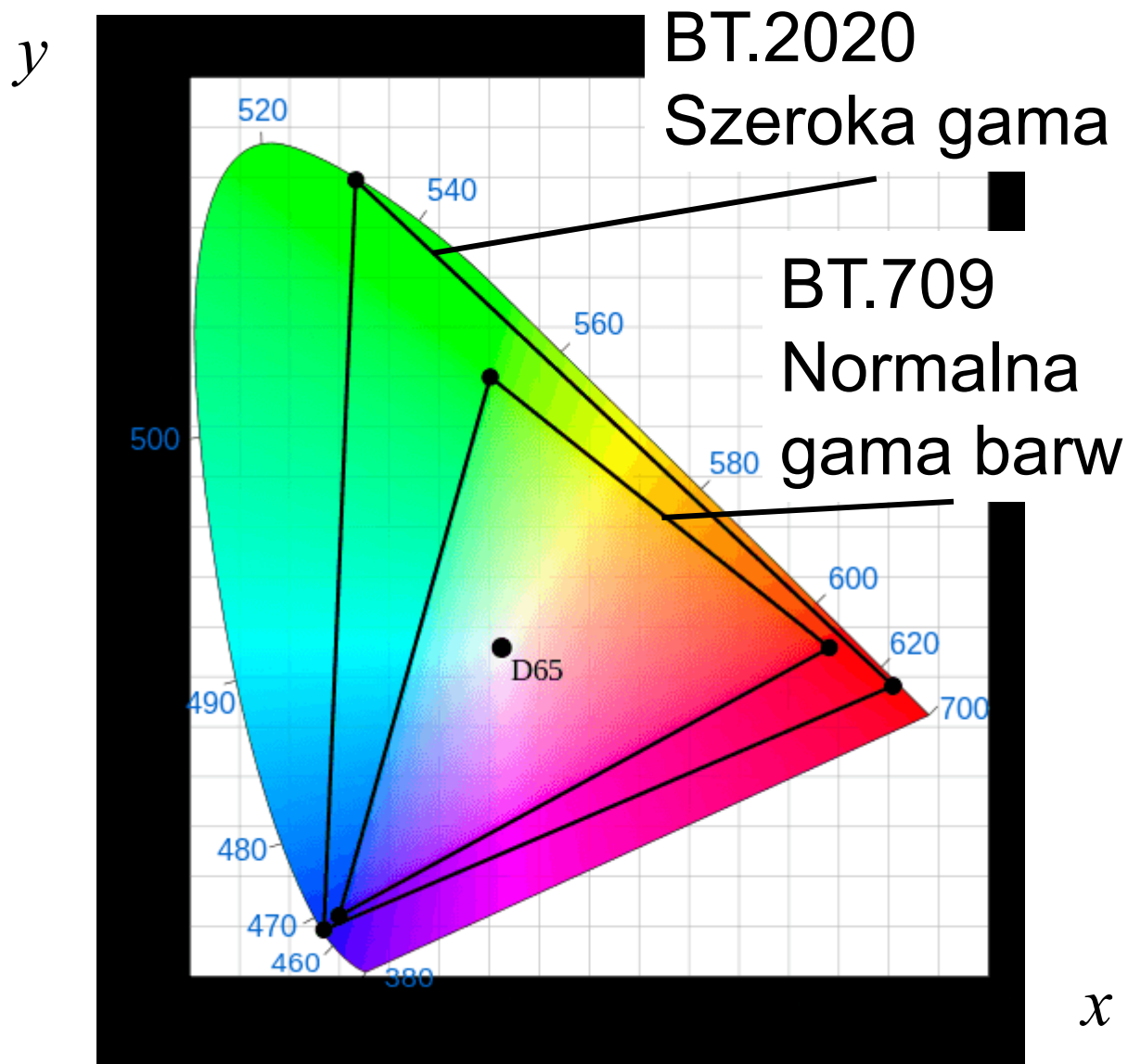
Gama barw



**Gama barw
monitora z
luminoforem
EBU**



Szeroka gama barw - Wide Color Gamut WCG



BT.2020

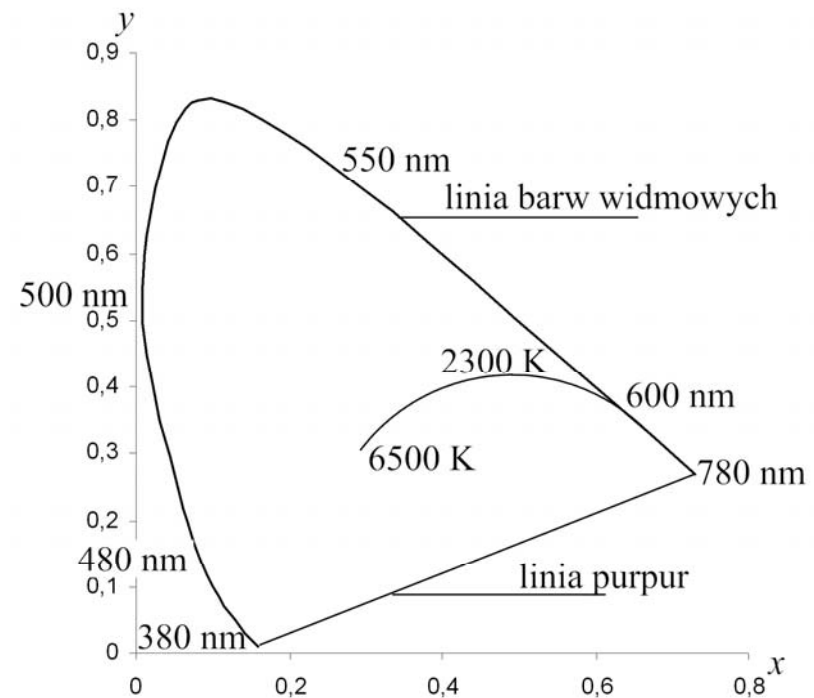
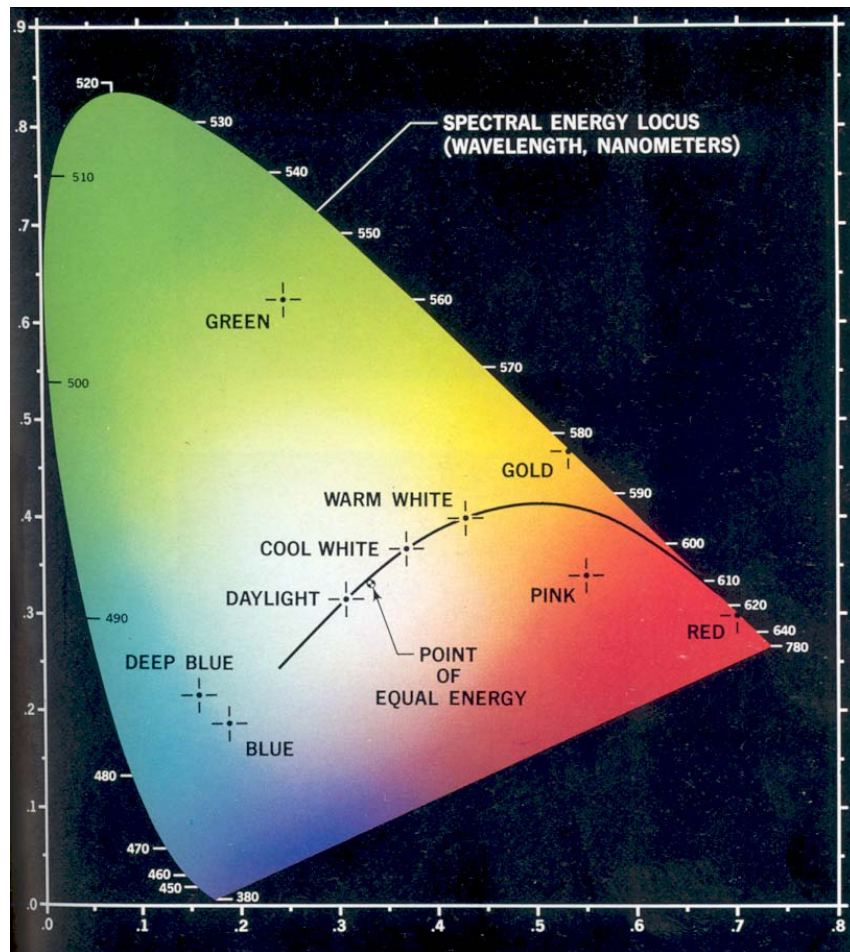
Widmowe barwy
podstawowe :

czzerwonej o długości fali
światłnej **630 nm**

(składowa **R**),
zielonej o długości fali
światłnej **532 nm**
(składowa **G**),

niebieskiej o długości
fali światłnej **467 nm**
(składowa **B**) .

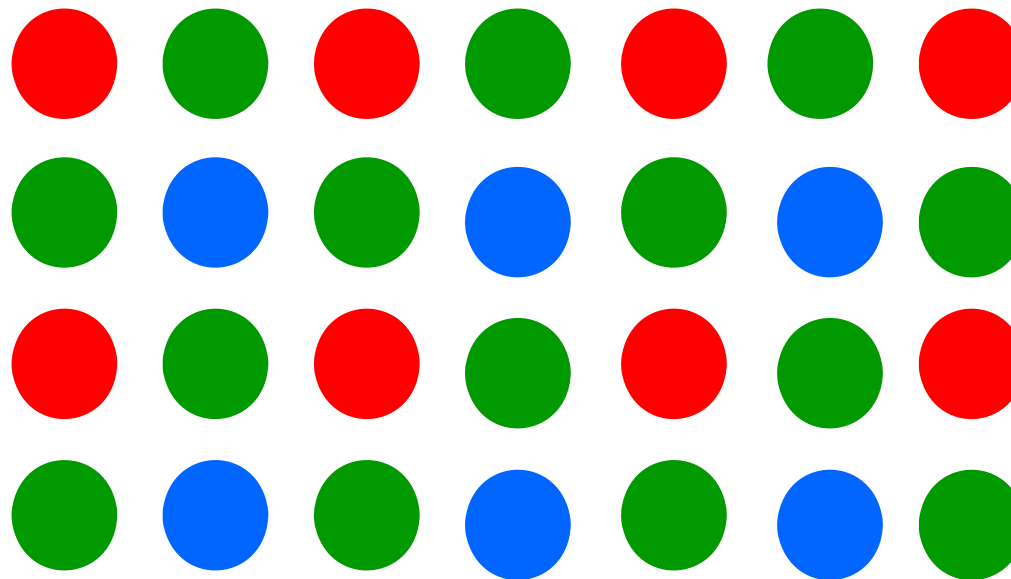
2.5.6. Temperatura barwowa



2.5.7. Akwizycja obrazów barwnych w kamerze

Zazwyczaj jest tylko jeden przetwornik

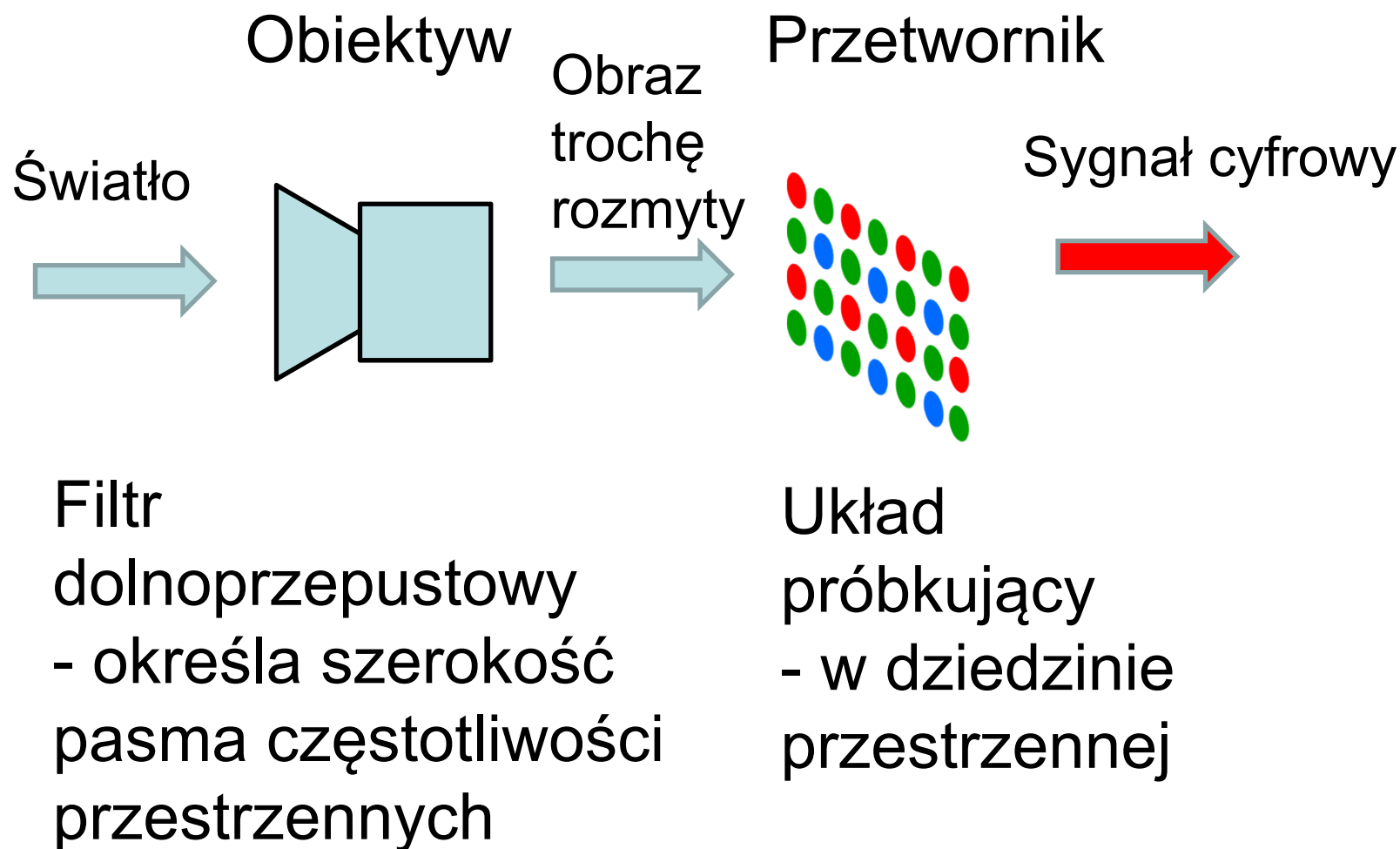
Elementy przetwornika przysłonięte filtrami



macierz

Bayera

Działanie kamery barwnej



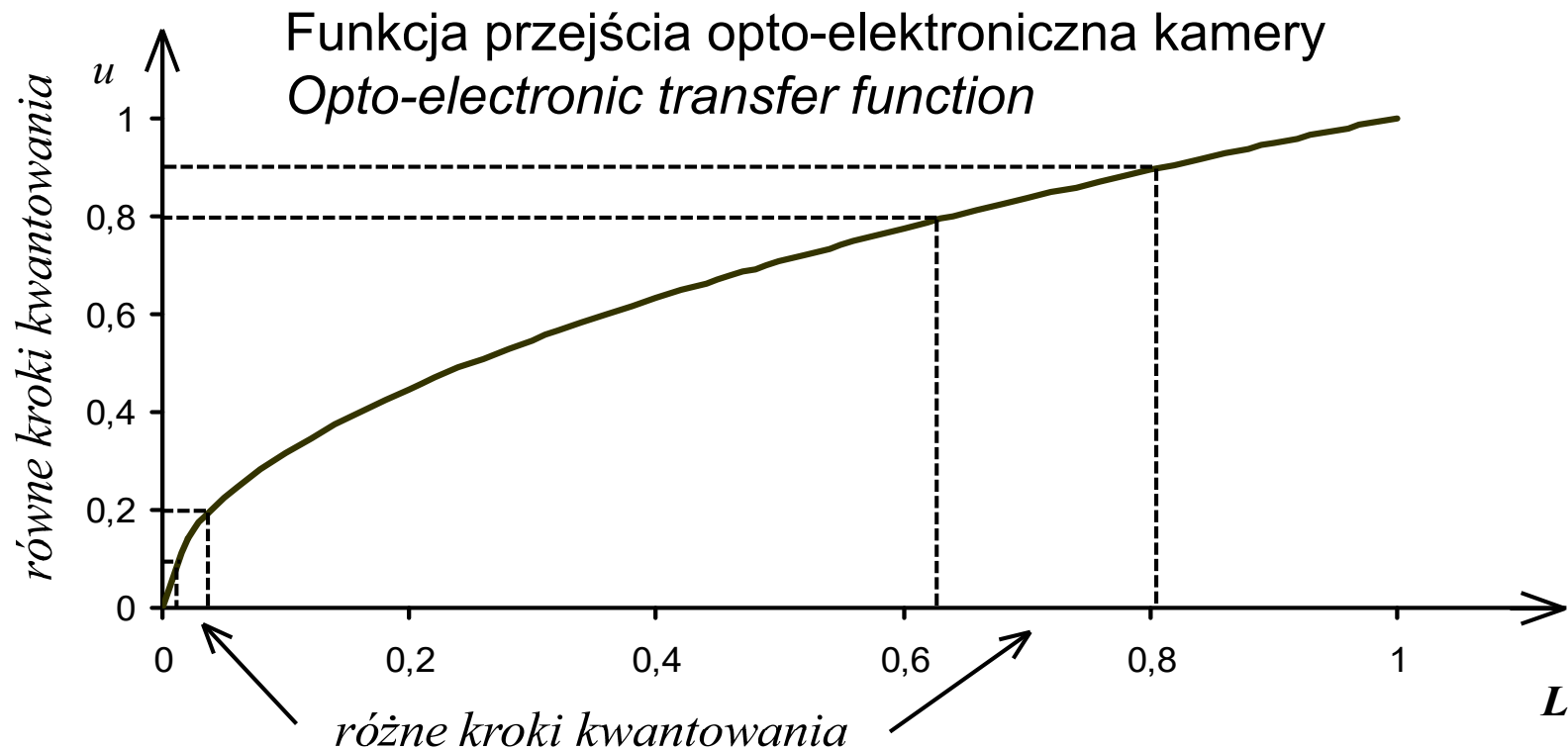
2.5.8. Funkcje przejścia: opto-elektroniczna i elektro-optyczna

Inne nazwy: Korekcja /transformacja gamma, Transformacja kontrastu

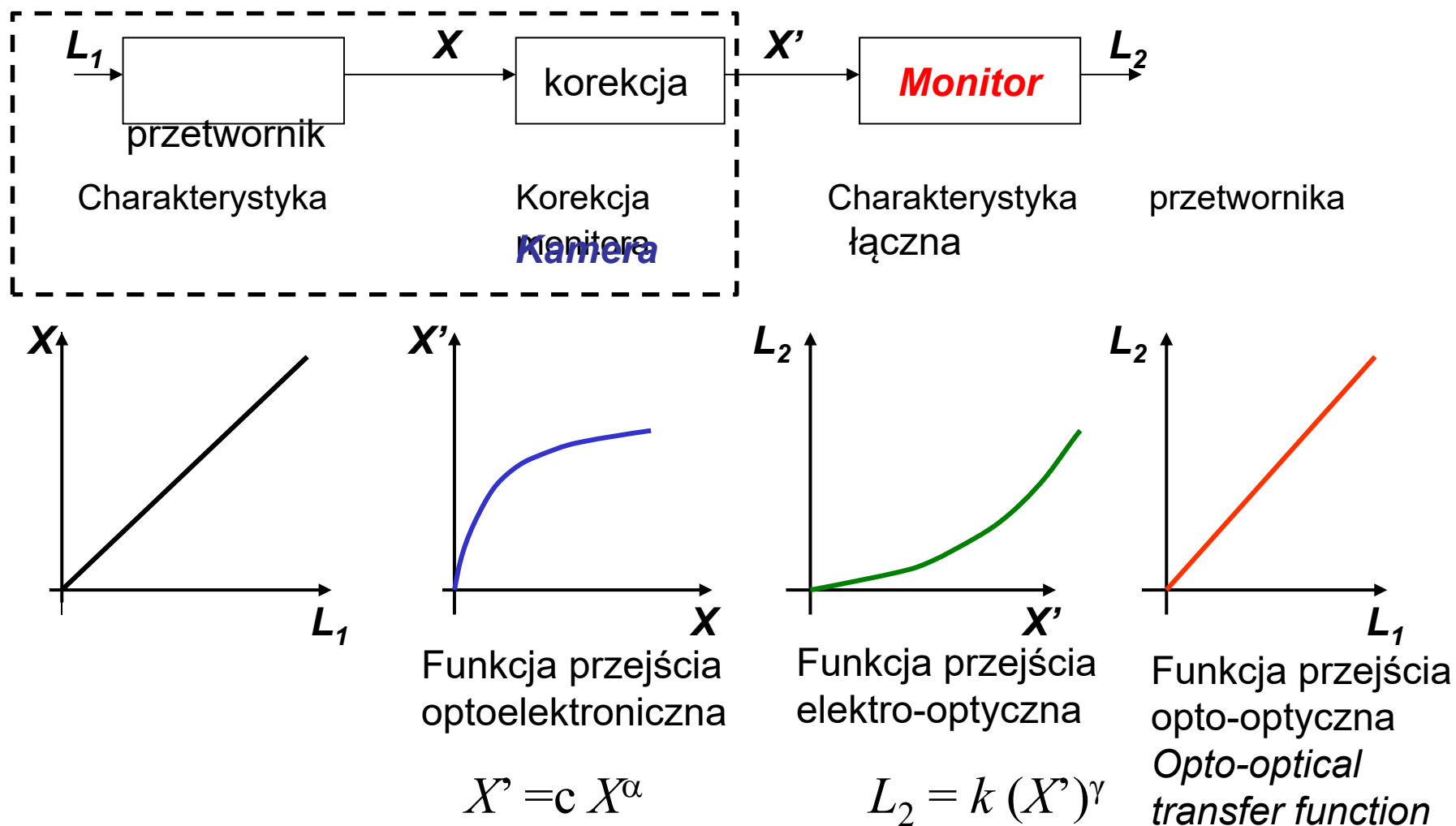
Prawo Webera-Fechnera

$$\frac{\Delta L}{L} \approx \text{const} \Rightarrow \text{Krok kwantowania} < \Delta L \approx \text{const} \cdot L$$

Kwantowanie równomierne po transformacji gamma



Funkcje przejścia: opto-elektroniczna i elektro-optyczna (c.d.)



Uwaga:

$$\alpha = 0,45$$

$$\gamma = 2,4$$

Powyższe zależności mają charakter uproszczony. Pominięto stałe addytywne i odkształcenia charakterystyk w pobliżu początku układu współrzędnych.

Model monitora

ITU-R zalecenie BT.1889

$$L = a \cdot (l+b)^\gamma$$

$\gamma = 2,4$, l – wartość składowej,

a – kontrast

b - jaskrawość

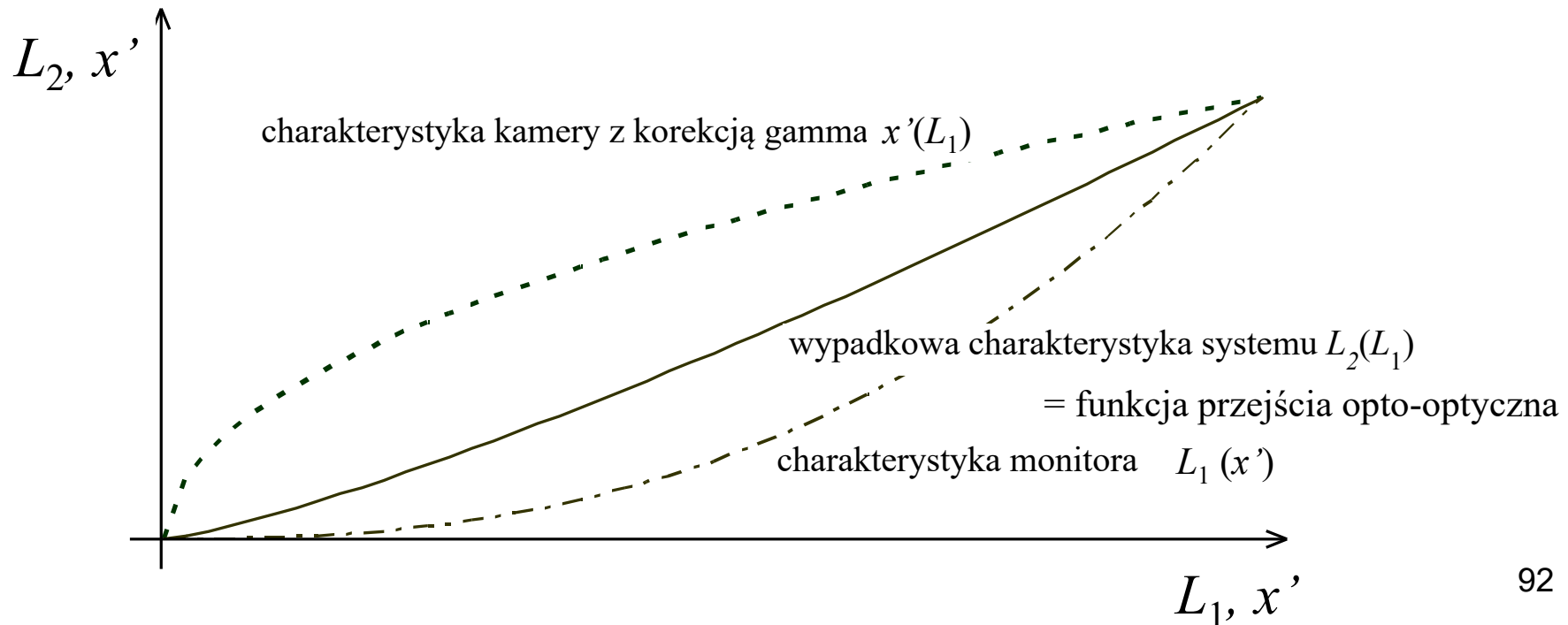
Transformacje składowych RGB (HDTV)

- ITU-R zalecenie BT.709 – telewizja HD

$$x' = \begin{cases} 4,5 x, & 0 \leq x < 0,018, \\ 1,099x^{0,45} - 0,099, & 0,018 \leq x \leq 1, \end{cases}$$

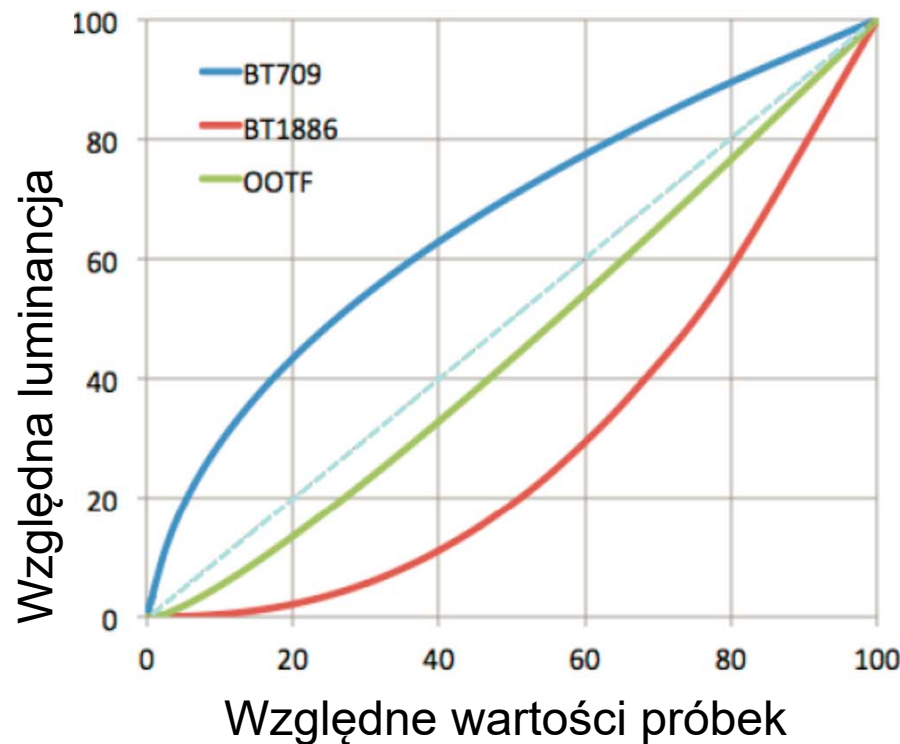
W przybliżeniu:

$$x' \approx x^{0,5} \Rightarrow L_2 \approx k(x')^{2,4}, \text{ to } L_2 \approx k(x')^{2,4} \approx k(x^{0,5})^{2,4} = kx^{1,20}$$



Wypadkowa funkcja przejścia całego systemu OOTF – Opto-Optical Transfer Function of the System

OOTF (System Gamma) in HD SDR



BT.709 (Camera)

Transfer Function
to Display $\approx 1/2.2$

Funkcja przejścia
opto-elektroniczna

BT.1886 (Display)

Display gamma for
flat panel display = 2.4
(2011)

Funkcja przejścia
elektro-optyczna

OOTF

Opto-Optical
Transfer Function
(System Gamma)
= 1.2

is known as good
picture in current TVs

Funkcja przejścia
opto-optyczna =
= Funkcja przejścia
wypadkowa =
= Funkcja przejścia
całego systemu

Współczesne systemy telewizyjne (SDR) mają
OOTF odpowiadającą wypadkowej wartości $\gamma > 1,0$

Funkcja przejścia opto-elektroniczna

Transformacja gamma (ciąg dalszy)

- sRGB (IEC 61699)
 - obraz nieruchomy głównie, urządzenia biurowe

$$u' = \begin{cases} 12,92u, & 0 \leq u < 0,0031308, \\ 1,055u^{\frac{1}{2,4}} - 0,055, & 0,0031308 \leq u \leq 1, \end{cases}$$

- UHDTV (ITU-R Rec. BT.2020) dla UHDTV (HDR)

$$u' = \begin{cases} 4,5u, & 0 \leq u < \beta, \\ \alpha u^{0,45} - (\alpha - 1), & \beta \leq u \leq 1, \end{cases}$$

$$\alpha = 1,099, \quad \beta = 0,018 \quad \text{dla próbek 10-bitowych,}$$

$$\alpha = 1,0993, \quad \beta = 0,0181 \quad \text{dla próbek 12-bitowych.}$$

2.5.9. Współrzędne transmisyjne barw

Właściwość: Składowe R , G i B są silnie skorelowane



Składowa czerwona



zielona



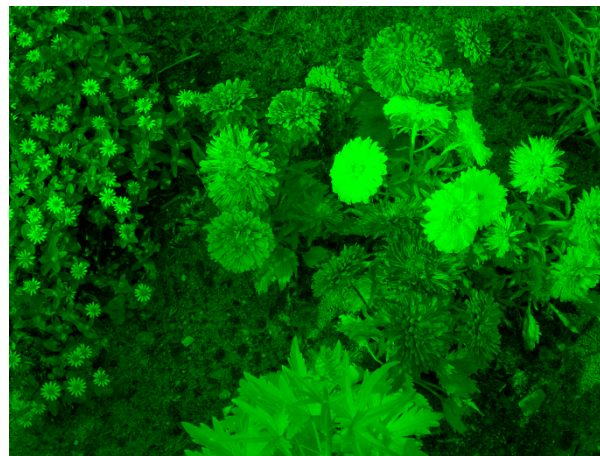
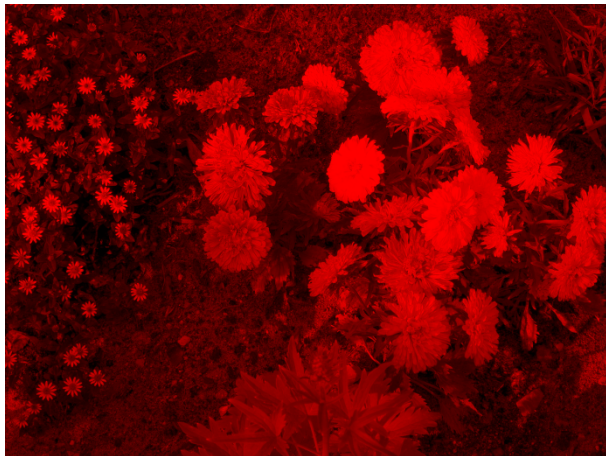
niebieska



Oryginalny obraz



Składowe R, G i B:



Luminancja - Luma

- **SDTV** (BT.601)

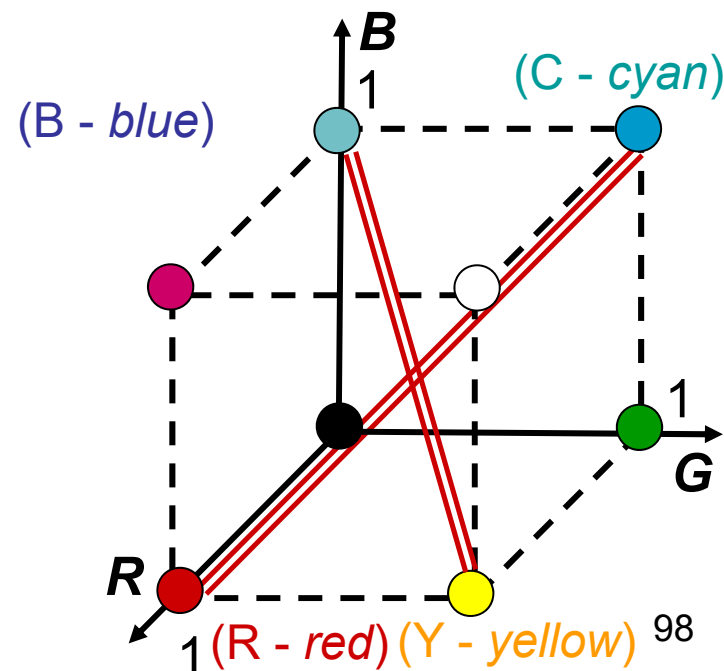
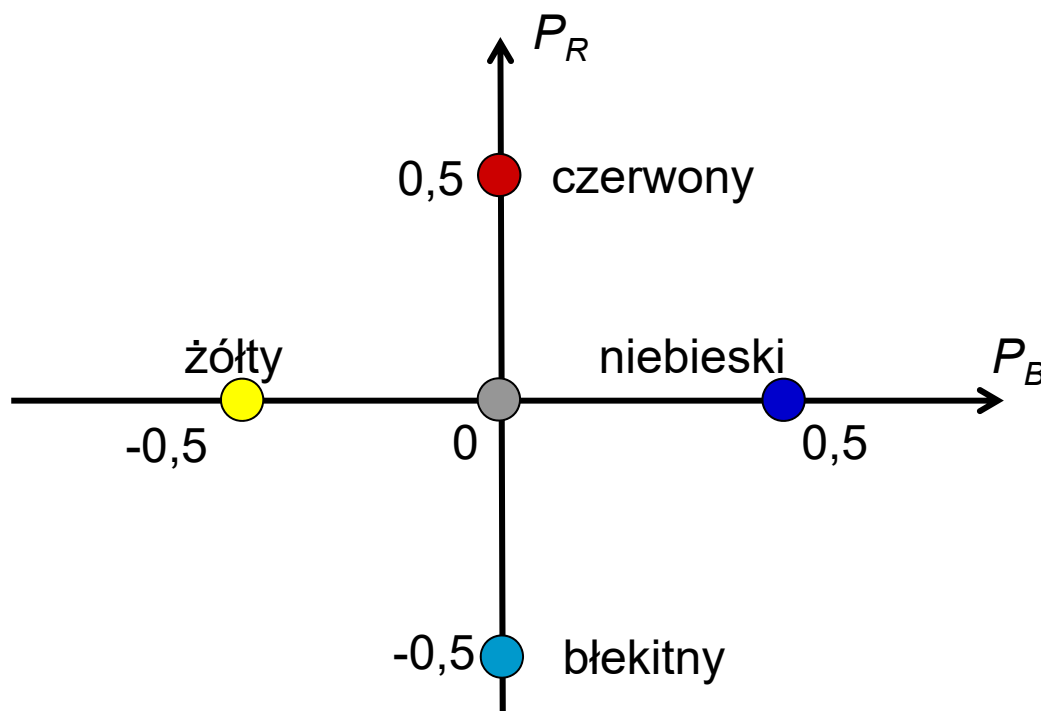
$$Y = 0,299 R' + 0,587 G' + 0,114 B'$$

- **HDTV** (BT.709)

$$Y = 0,2126 R' + 0,7152 G' + 0,0722 B'$$

Chrominancja - Chroma

- $P_B = 0,564 (B' - Y)$
 $= -0,169 R' - 0,331 G' + 0,500 B'$,
Według BT.601
- $P_R = 0,713 (R' - Y)$
 $= 0,500 R' - 0,419 G' - 0,081 B'$,
- Chromy P_B i P_R reprezentują barwę w stosunku do par barw przeciwstawnych "czerwona-błękitna" (P_R) oraz "niebieska-żółta" (P_B)



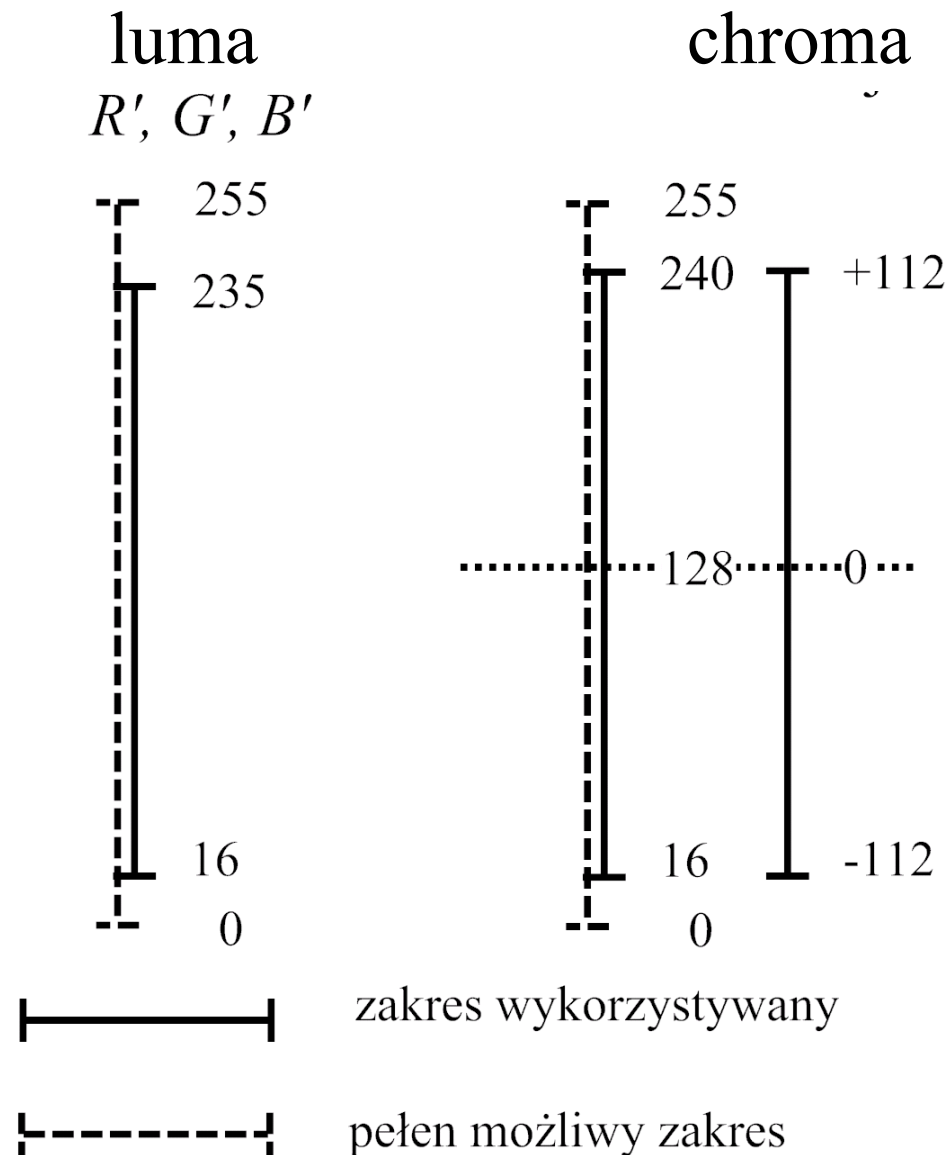
Chrominancja / chroma cyfrowa

Określone specjalnie dla cyfrowych
sygnałów wizyjnych składowe chromy
 C_B , C_R definiuje się jako np. wg BT.601

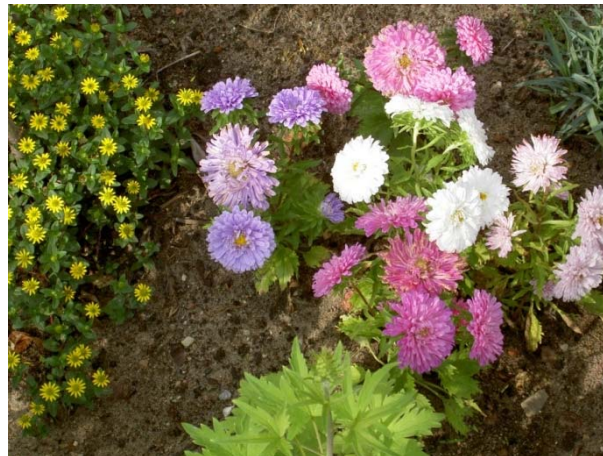
$$\begin{aligned} C_B &= 0,564(B' - Y) \\ &= -0,169R' - 0,331G' + 0,500B', \end{aligned}$$

$$\begin{aligned} C_R &= 0,713(R' - Y) \\ &= 0,500R' - 0,419G' - 0,081B'. \end{aligned}$$

Sygnały wizyjne



Oryginał



Luma



Chroma C_B



Chroma C_R



Luminancja i chrominancje

Luma i chromy

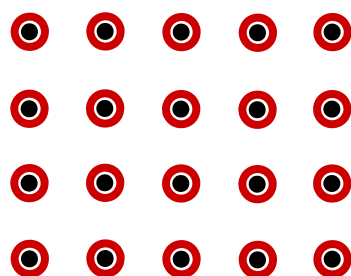


Próbkowanie chromy

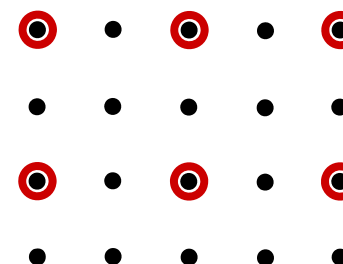
- próbka lummy

○ próbka chromy

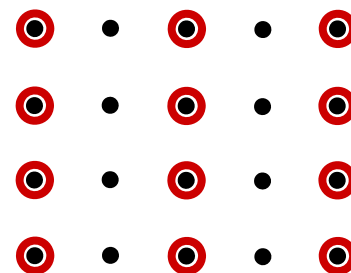
4:4:4



4:2:0



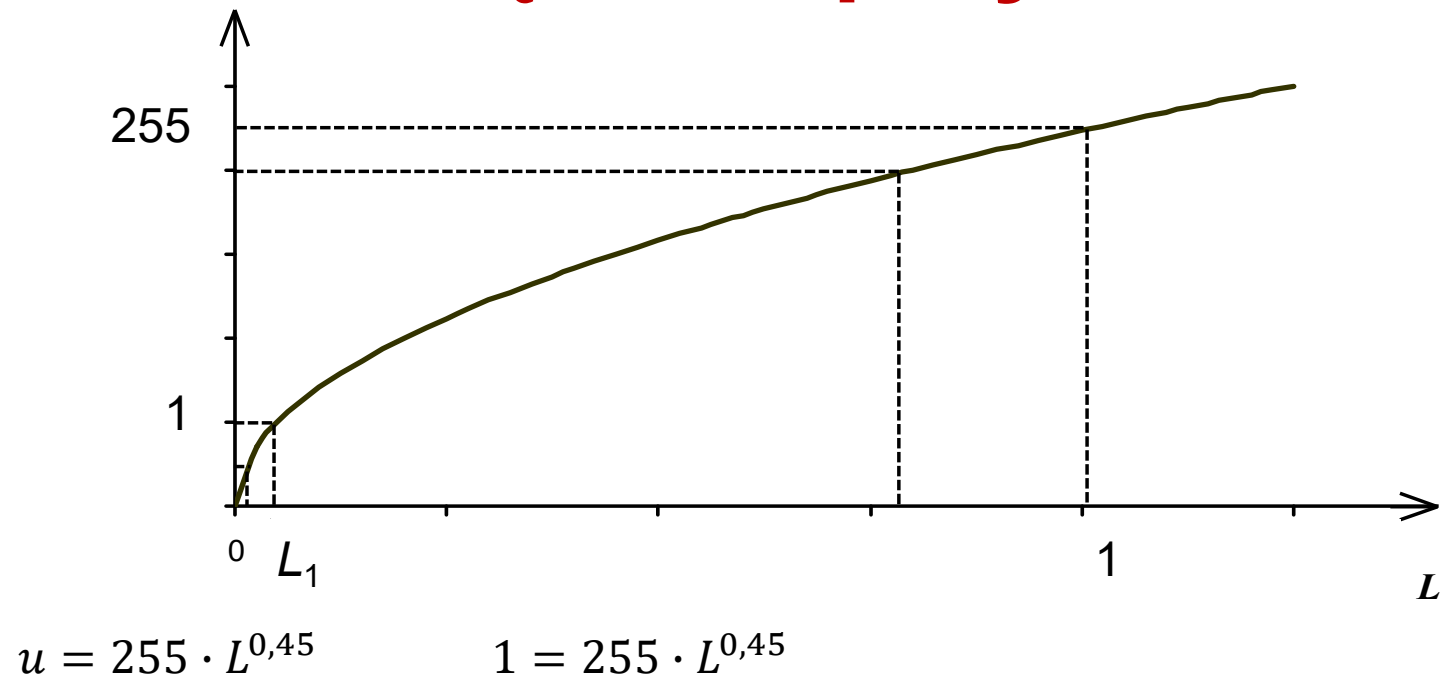
4:2:2



Przykład

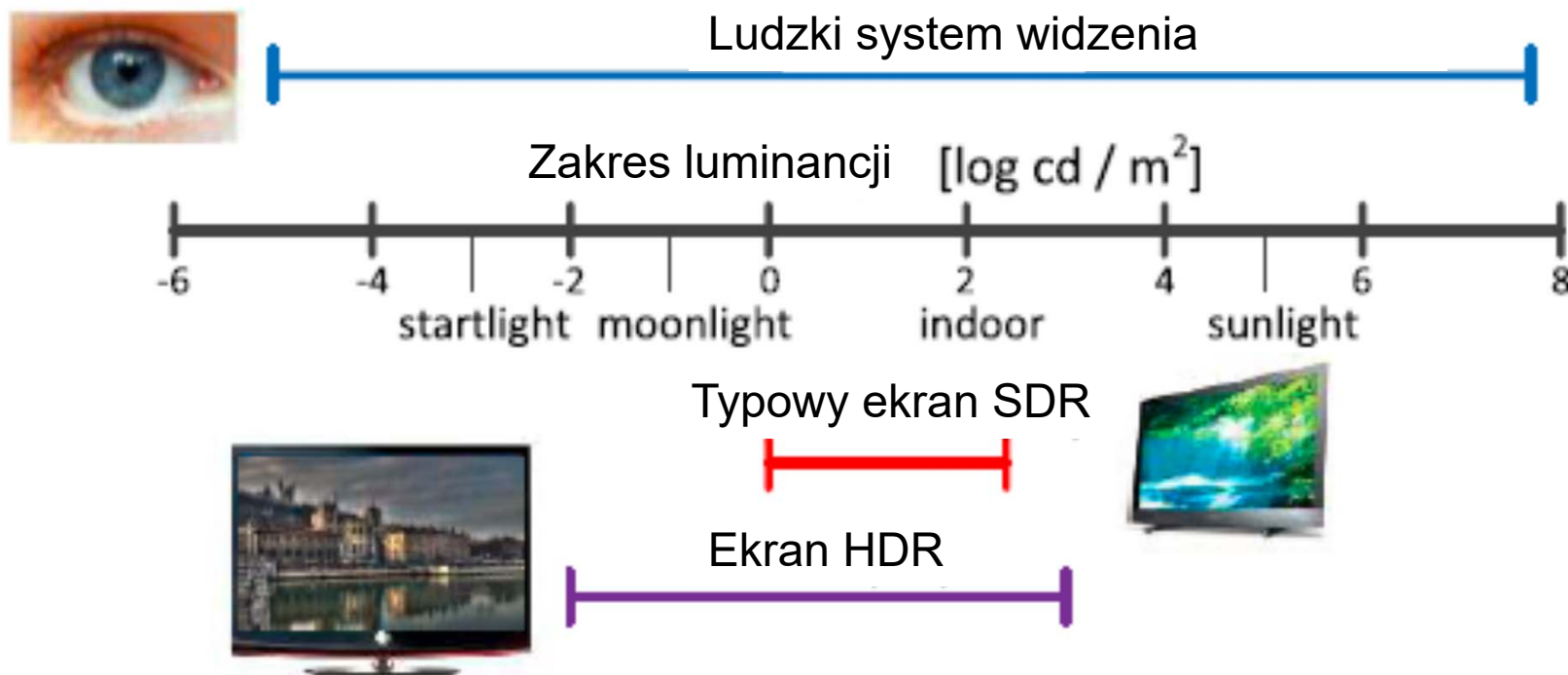
- Określ słownie barwę światła reprezentowanego przez sygnał :
- **1)** $Y = 200$, $C_R = 245$, $C_B = 20$ (wartości reprezentacji ośmiobitowych bez znaku);
- **2)** $Y = 180$, $C_R = 125$, $C_B = 130$ (wartości reprezentacji ośmiobitowych bez znaku).

Rozwiązanie przykładu 2



$$222\,772 \leq 2^{17}$$

2.6. Duży zakres dynamiczny High Dynamic Range HDR



Duży zakres dynamiczny

High Dynamic Range HDR

Ekrany dla zakresu dynamicznego:

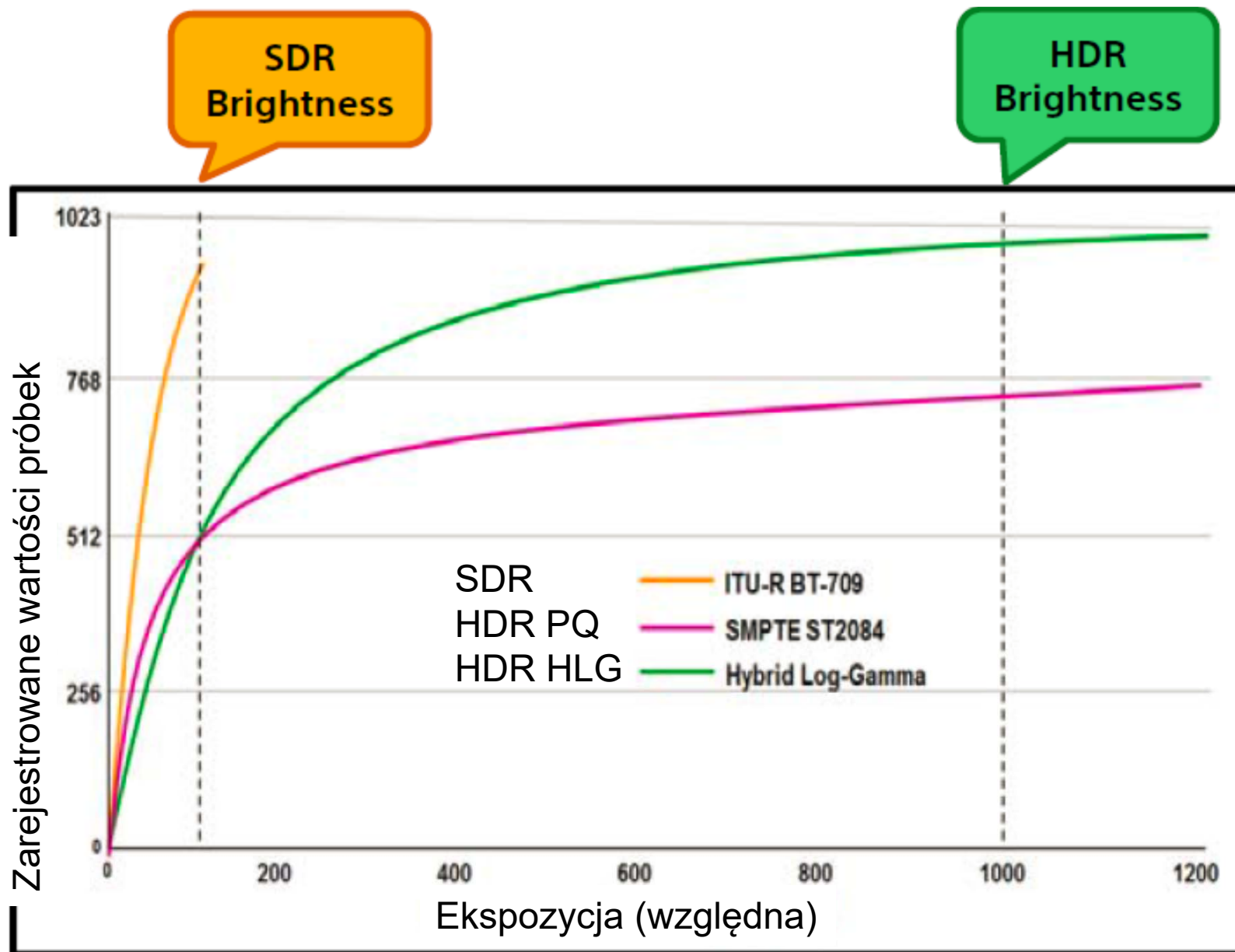
SDR (Standard Dynamic Range) : 0.1 – 100 nit (cd/m²)

HDR (High Dynamic Range) : 0.0005 – 10 000 nit (cd/m²)

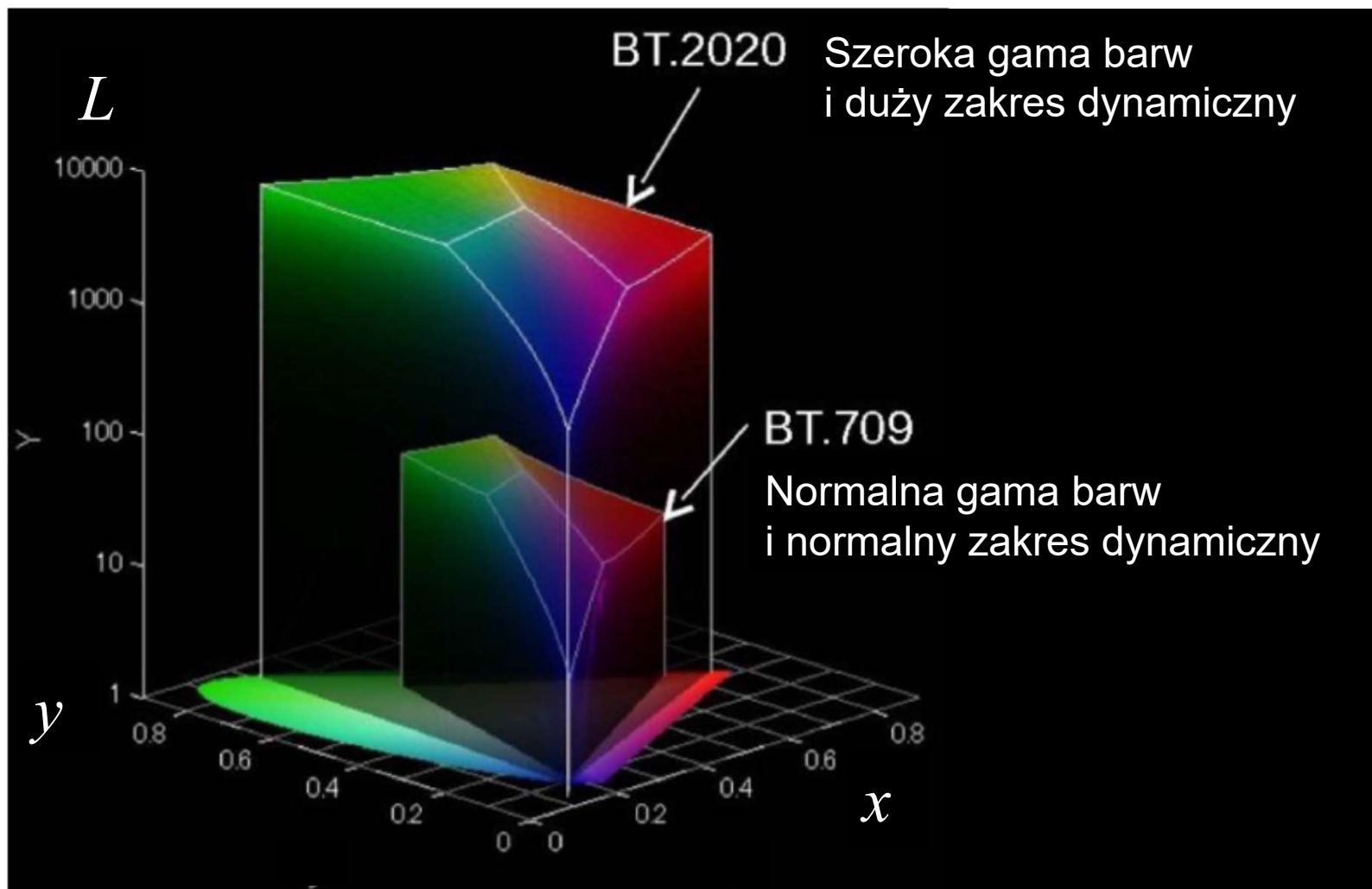
Dwa typy funkcji opto-elektronicznej OETF
(**10 bit/sample**) :

- Perceptual Quantizer (PQ)
- Hybrid Log-Gamma (HLG)

HDR OETF



Szeroka gama barw - Wide Color Gamut WCG



Przykład 1

- Wartości składowych R , G , B są określone liczbami 8-bitowymi.
 $R = 125$, $G = 230$, $B = 1280$.
Wyznaczyć 8-bitowe wartości składowych C , M , Y .
- Wyznaczyć składowe transportowe.

Przykład

- Próbkі sygnału elektrycznego reprezentującego sygnał świetlny są wyrażane za pomocą liczb 8-bitowych o poziomach wyznaczonych z funkcji wykładniczej o wykładniku $\alpha = 0,45$.
- Ile bitów reprezentacji próbek potrzeba, by sygnał równomiernie skwantowany miał kroki kwantowania nie dłuższe niż w systemie opisanym wyżej.

2.7. Potok przesyłania wizji



RGB

Kamera

Poprawa jakości obrazu

Kompresja (kodowanie źródłowe)

Szyfrowanie obrazu

Kodowanie kanałowe

Enkapsulacja – Transmisja

Dekodowanie kanałowe

Deszyfracja

Dekompresja (dekodowanie źródłowe)

Przetwarzanie obrazu

Przetwarzar

Monitor

Projektor

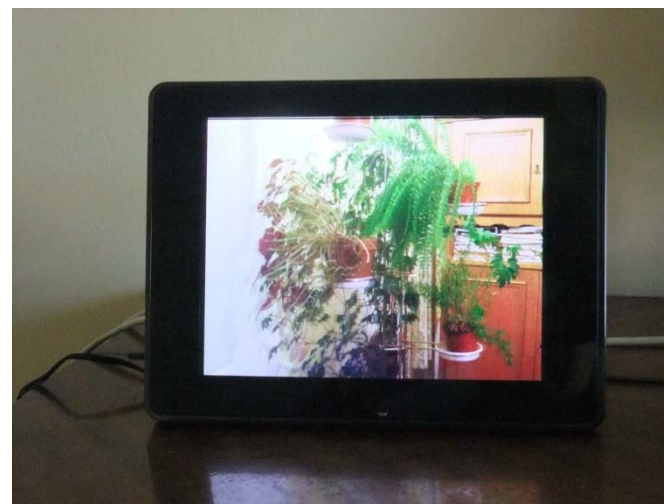
2.8. Wprowadzenie do obrazów stereoskopowych

Stereofotografia

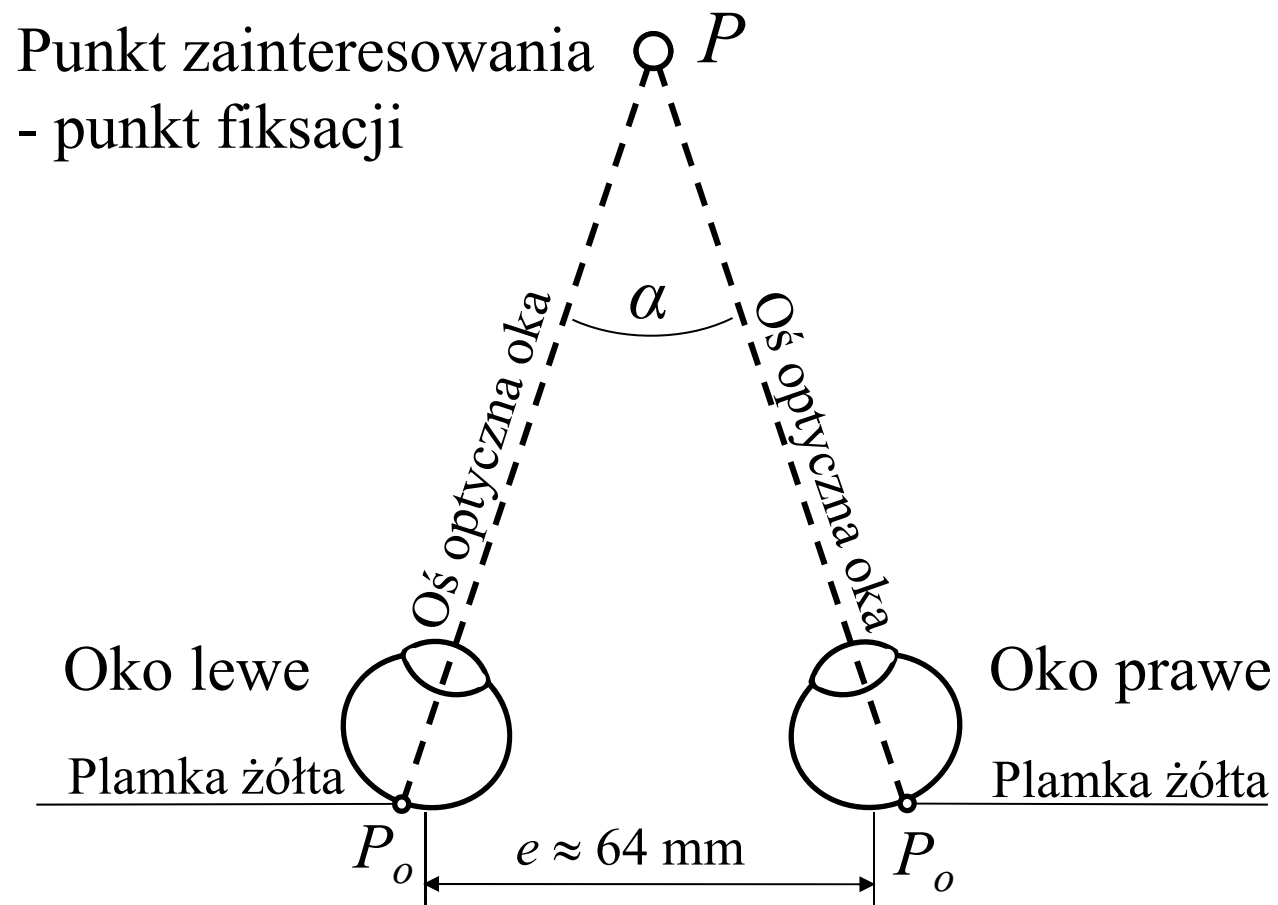
Przed ~ 100 laty



Dzisiaj

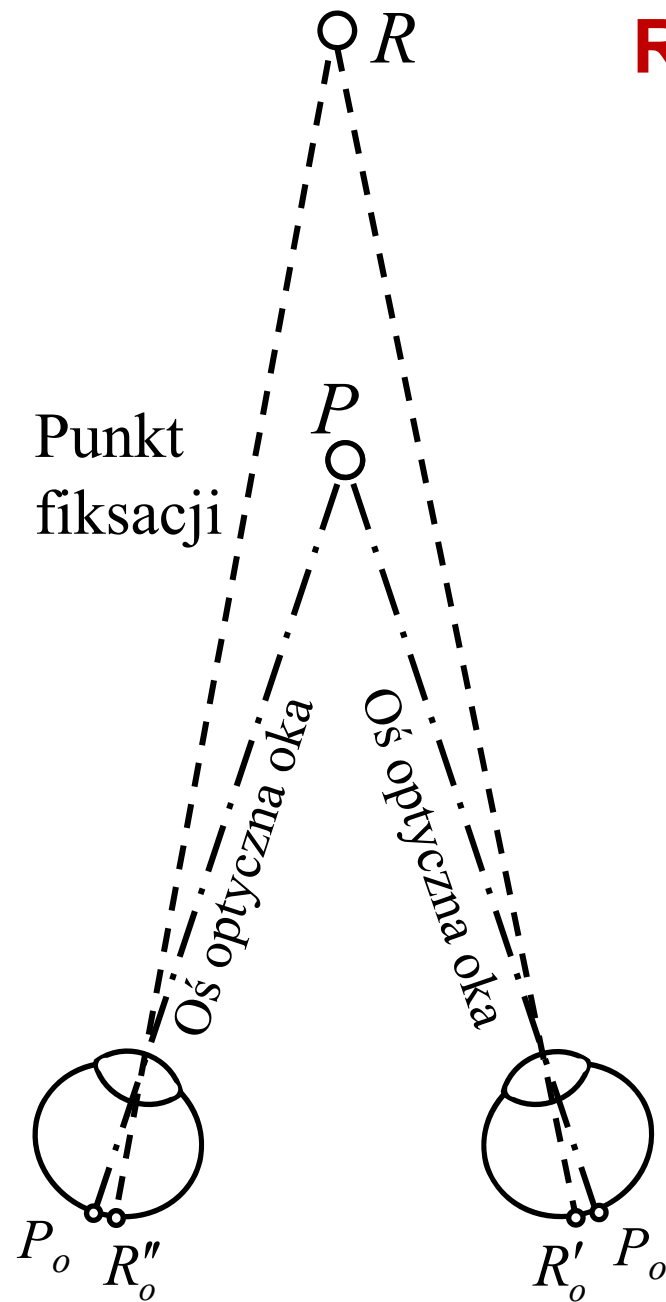


Paralaksa i zbieżność

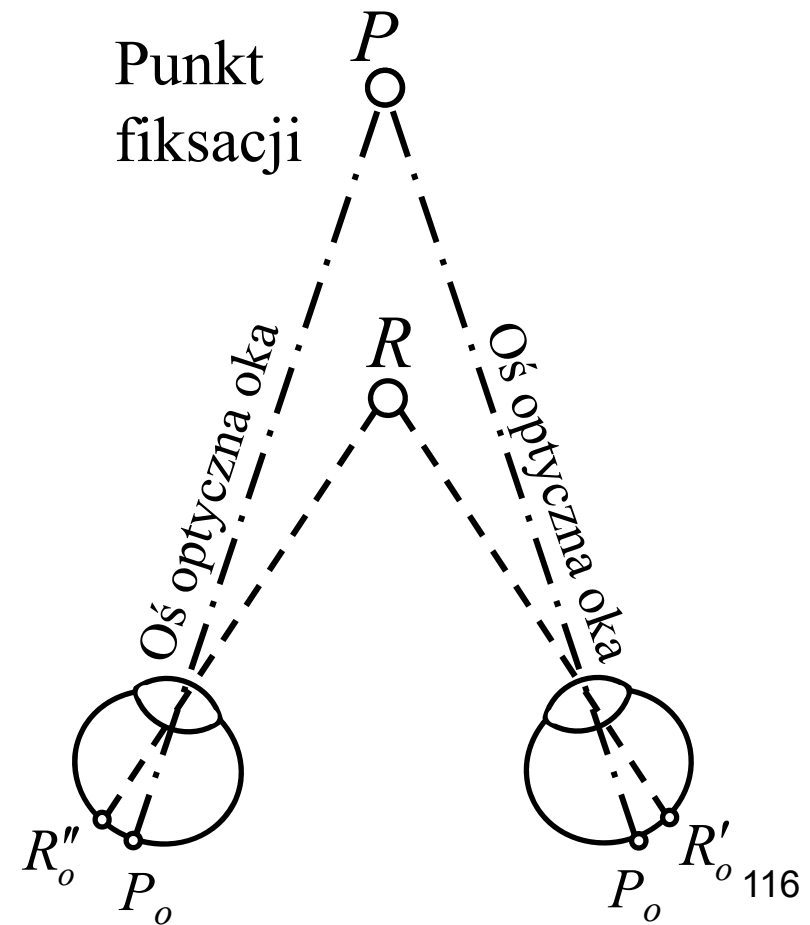


Rozbieżność siatkówkowa

a)



b)



Jednooczne wskazówki głębi

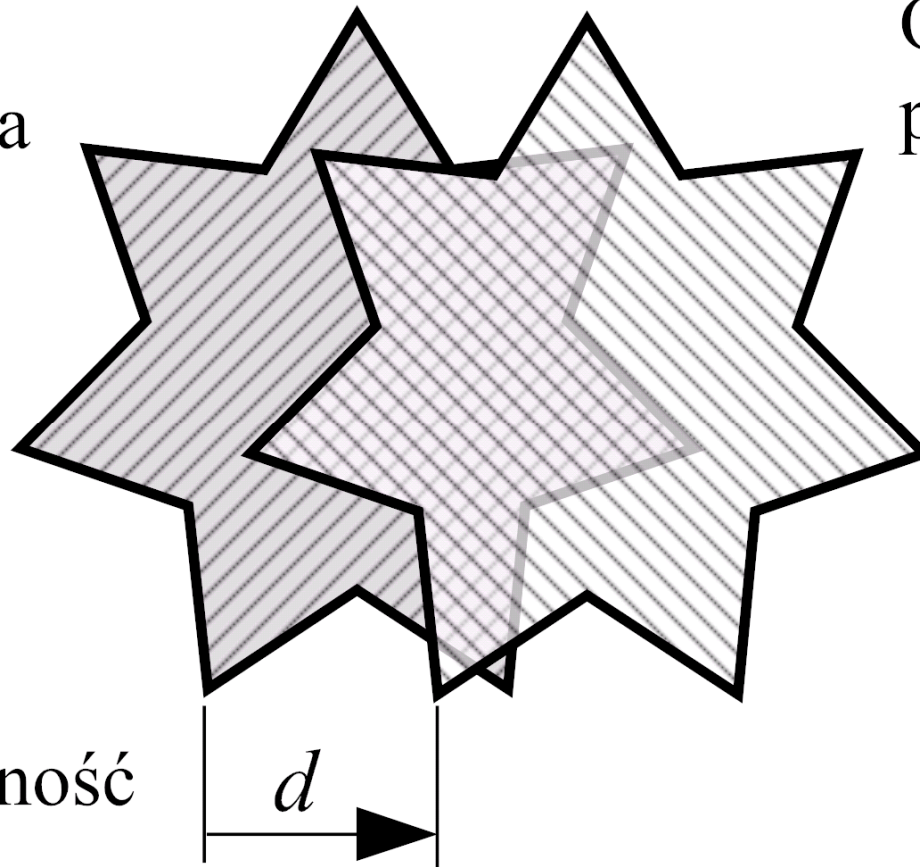
- **Akomodacja**
- **Przesłanianie**
- **Względny rozmiar**
- **Perspektywa, oświetlenie i cienie**
- **Gradient tekstury**
- **Paralaksa ruchu**
- **Zamglenie, rozmycie i zmniejszenie kontrastu**

Rozbieżność (*disparity*)

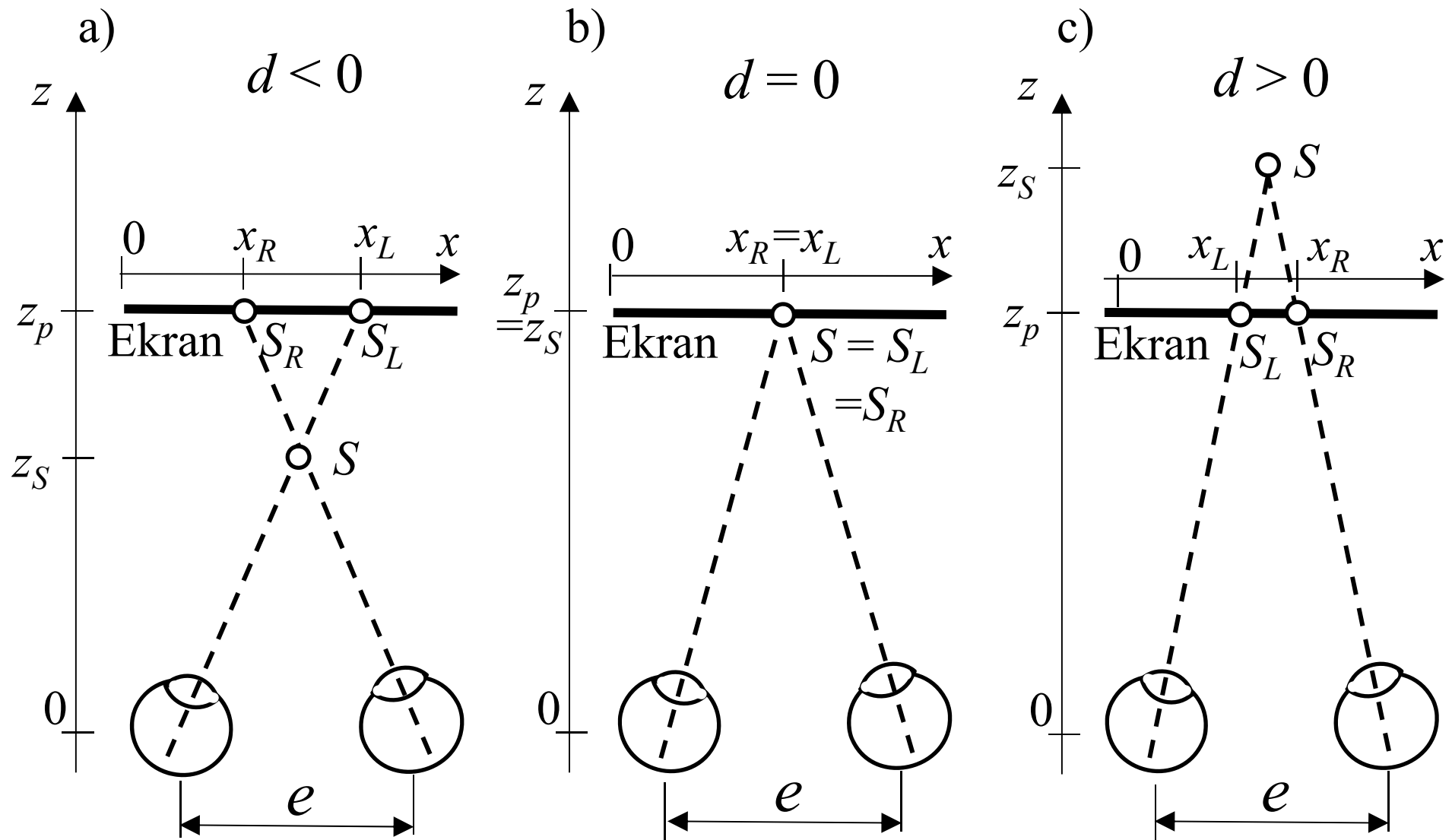
Obraz dla
lewego oka

Obraz dla
prawego oka

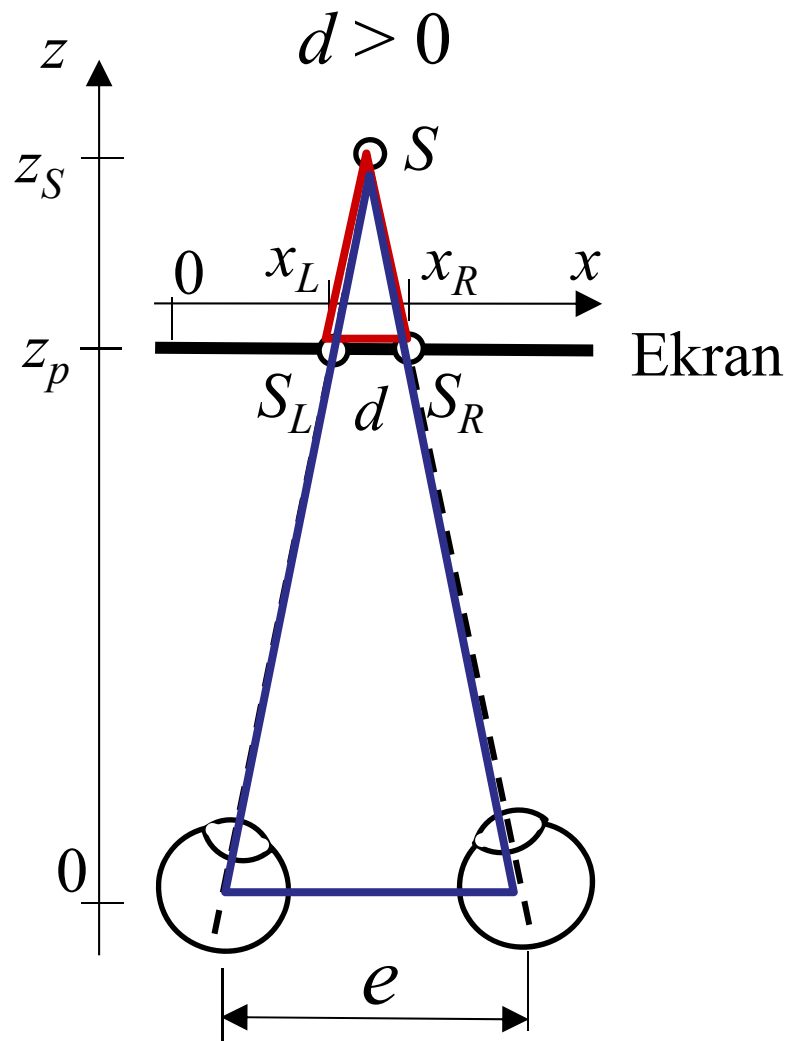
Rozbieżność



Prezentacja obrazu stereoskopowego na płaskim ekranie

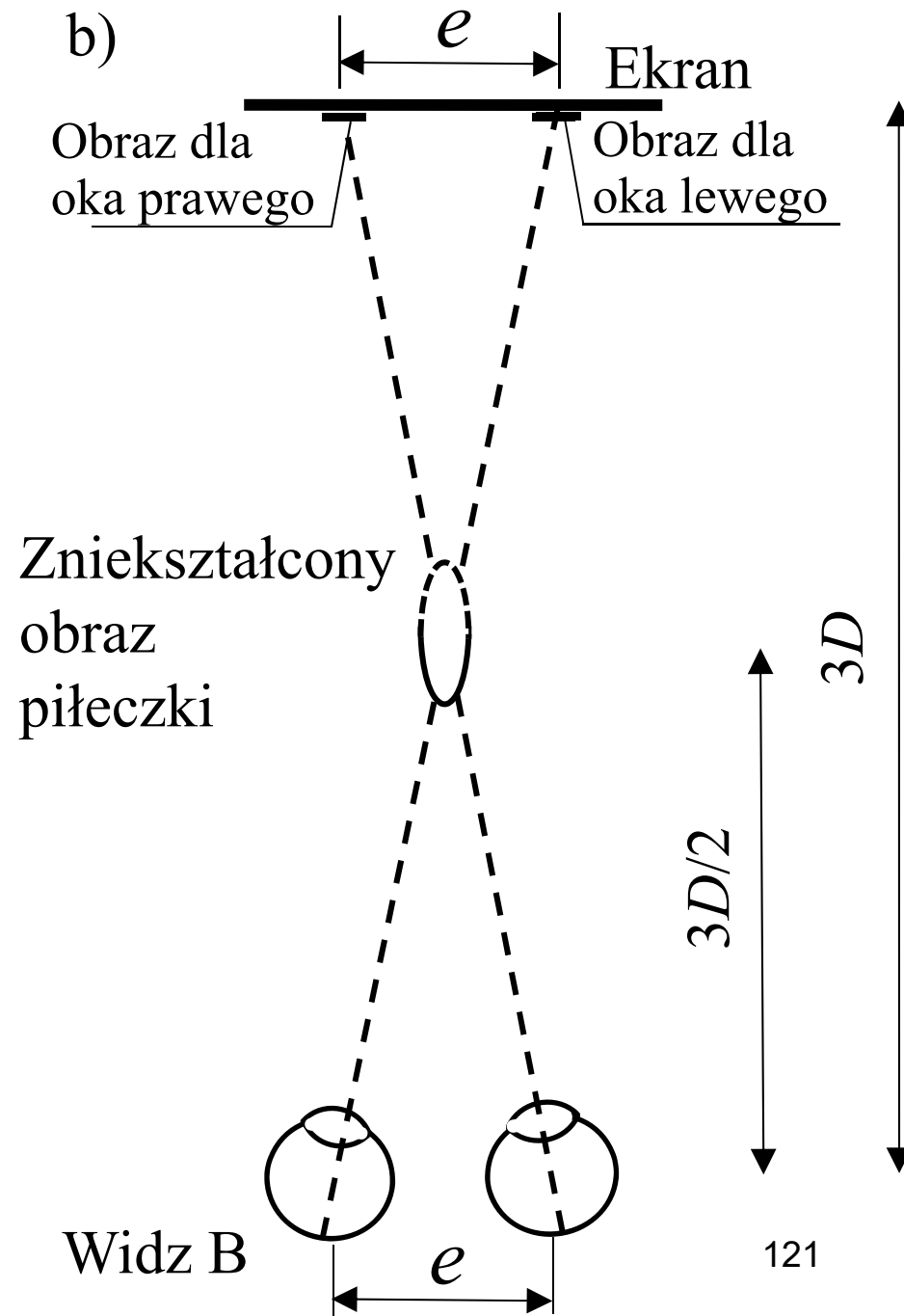
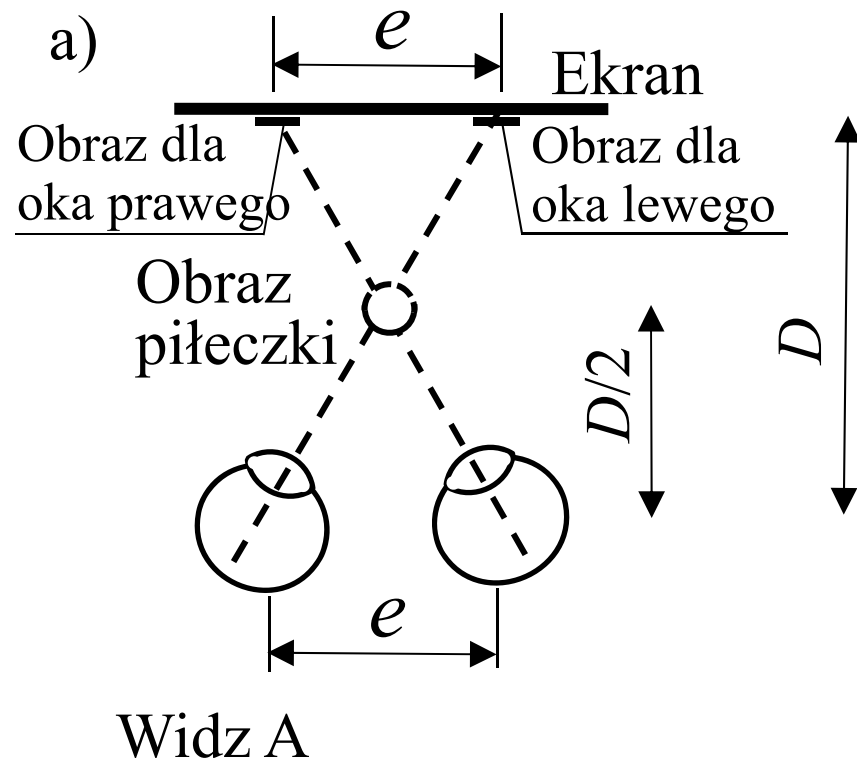


Prezentacja obrazu stereoskopowego na płaskim ekranie



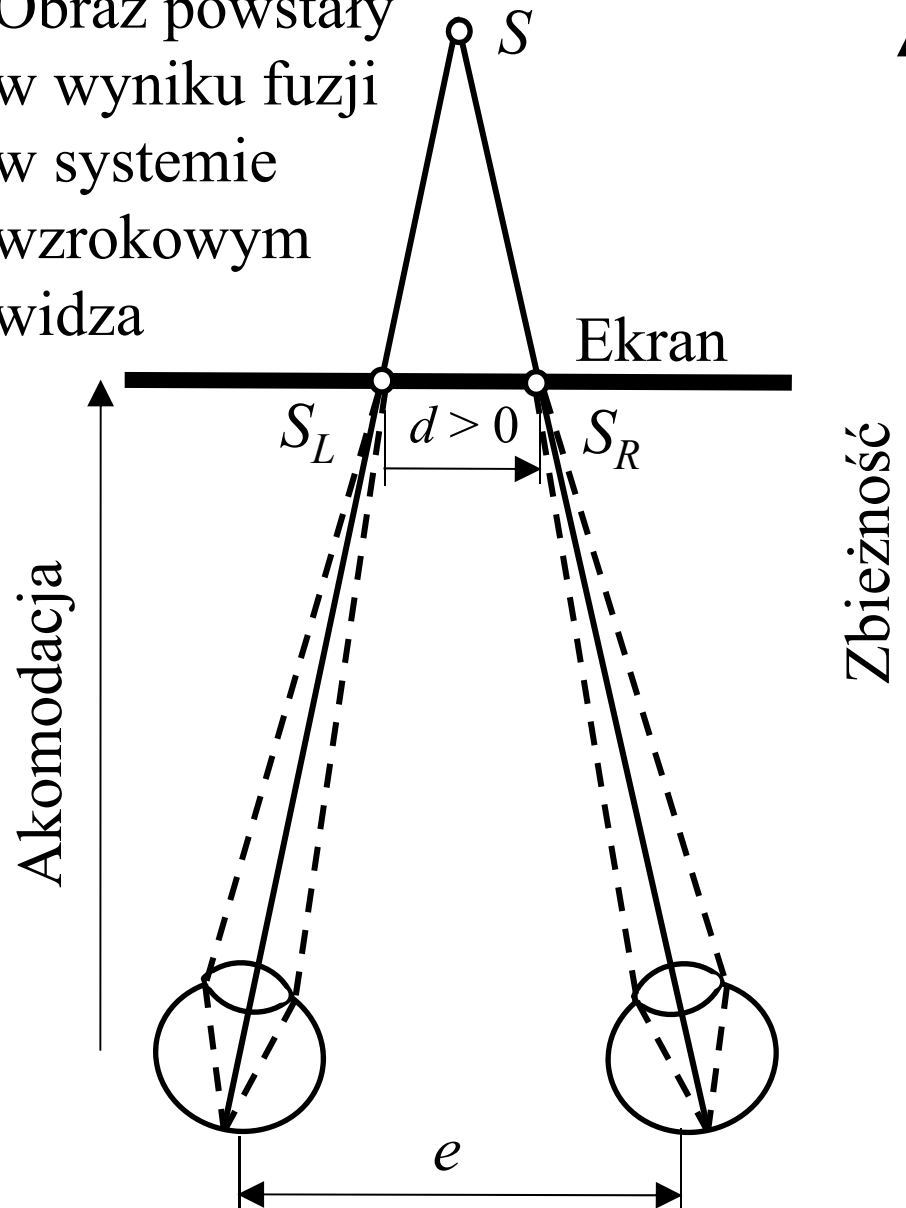
$$\frac{z_s - z_p}{d} = \frac{z_s}{e}$$

$$z_s = \frac{z_p \cdot e}{e - d}$$

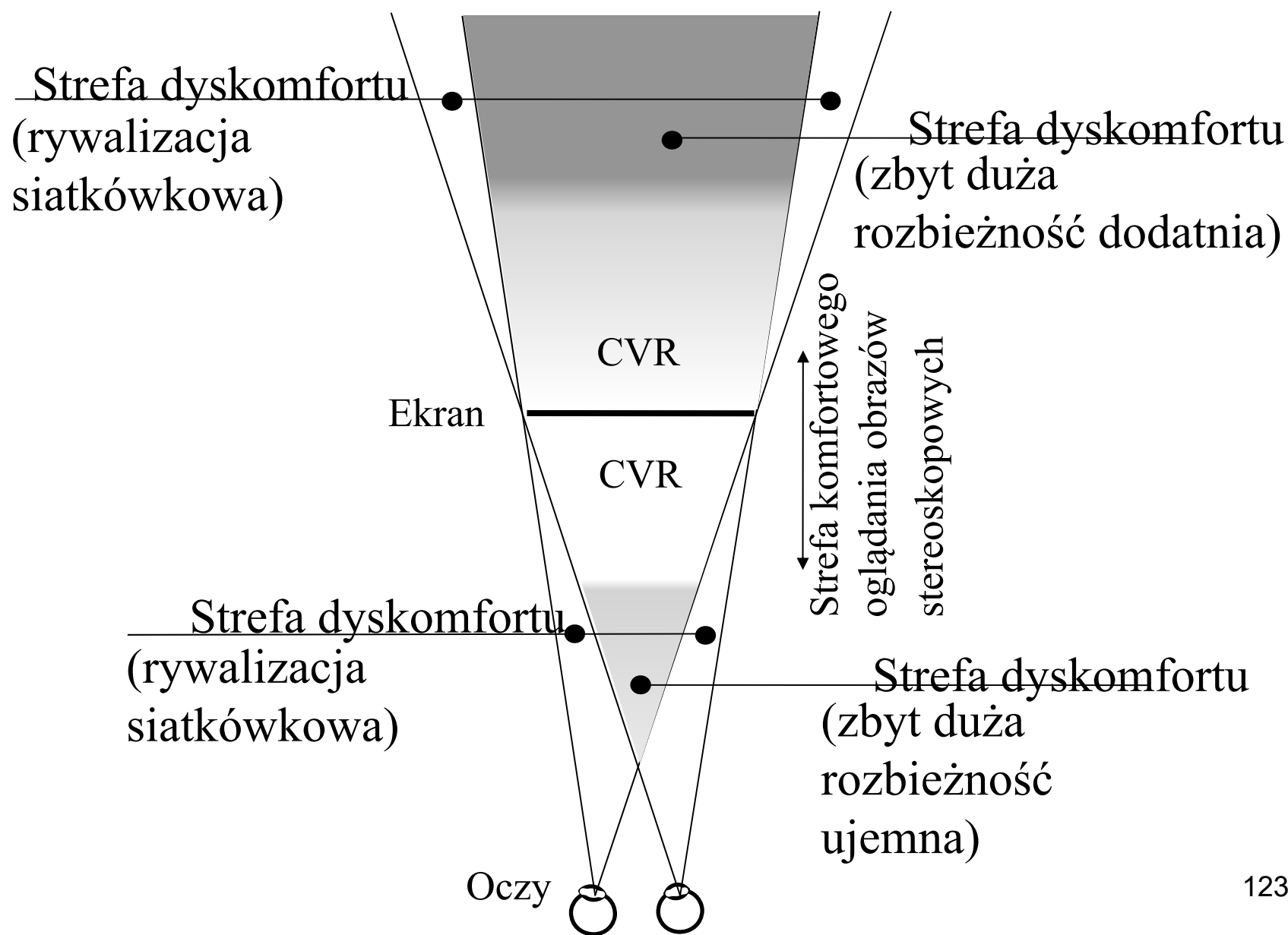


Konflikt pomiędzy akomodacją i zbieżnością

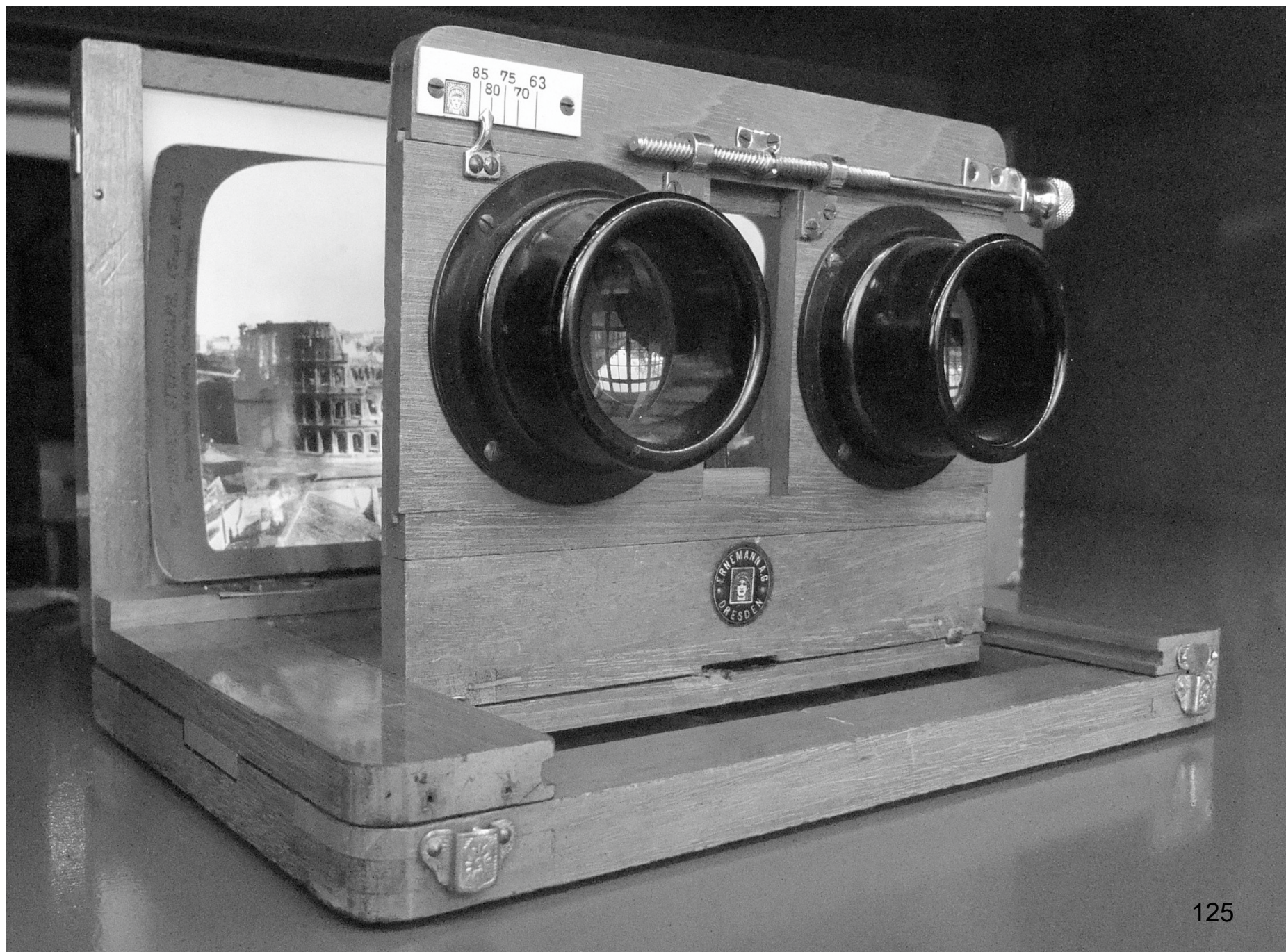
Obraz powstały
w wyniku fuzji
w systemie
wzrokowym
widza



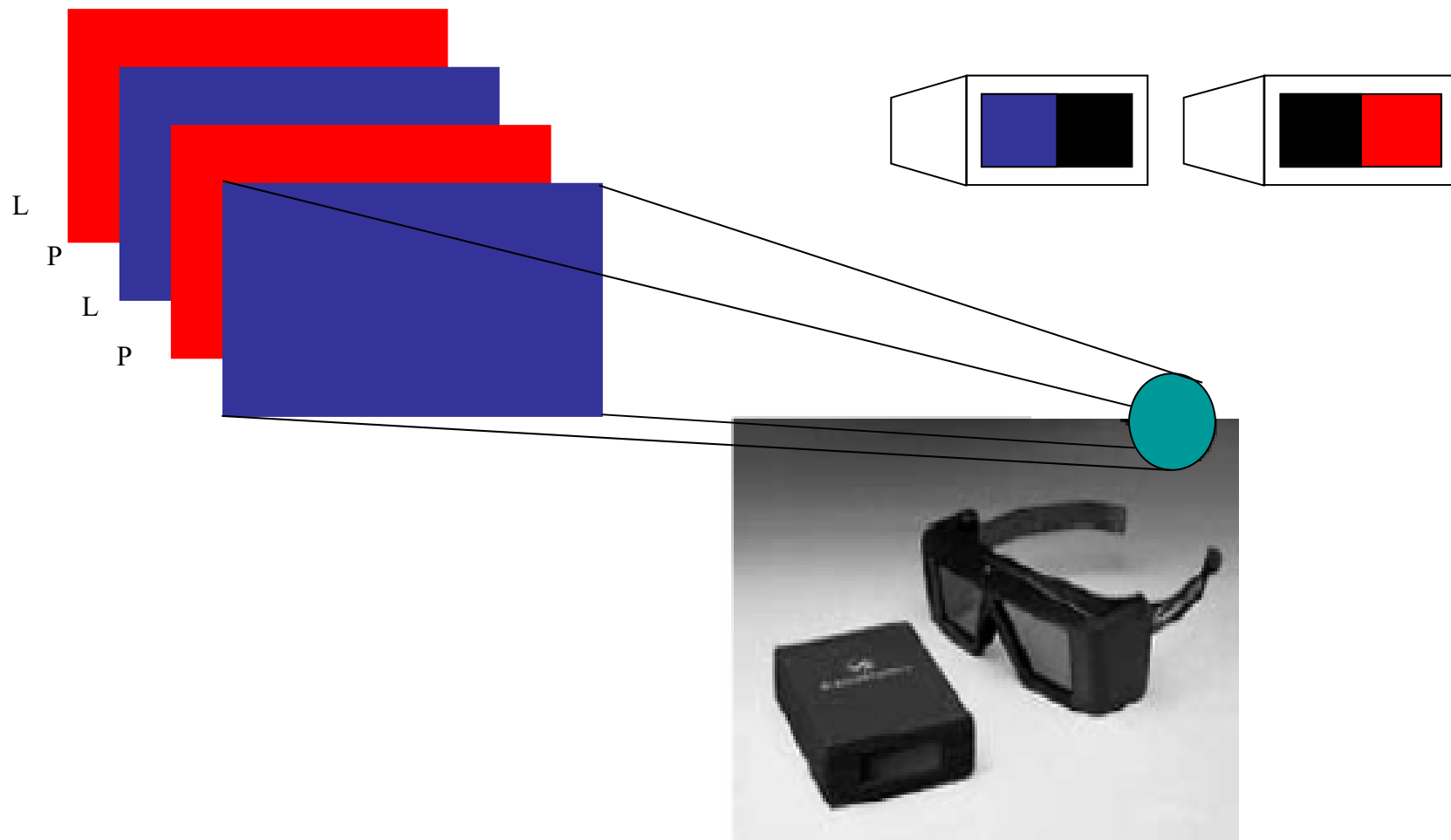
Konflikt pomiędzy akomodacją i zbieżnością



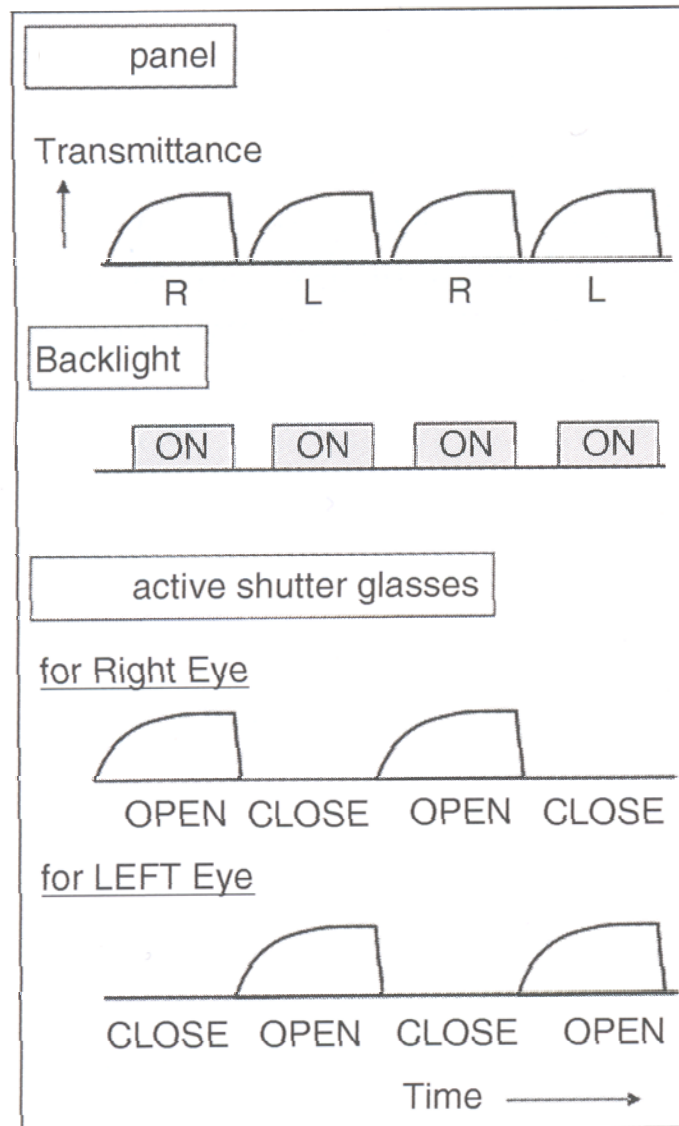
Separacja obrazów



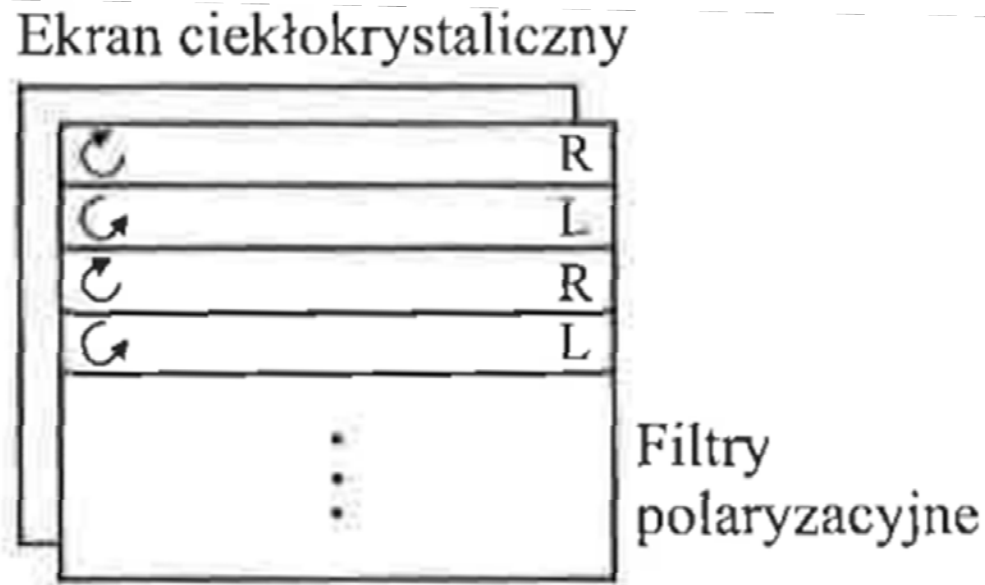
Wyświetlacze migawkowe



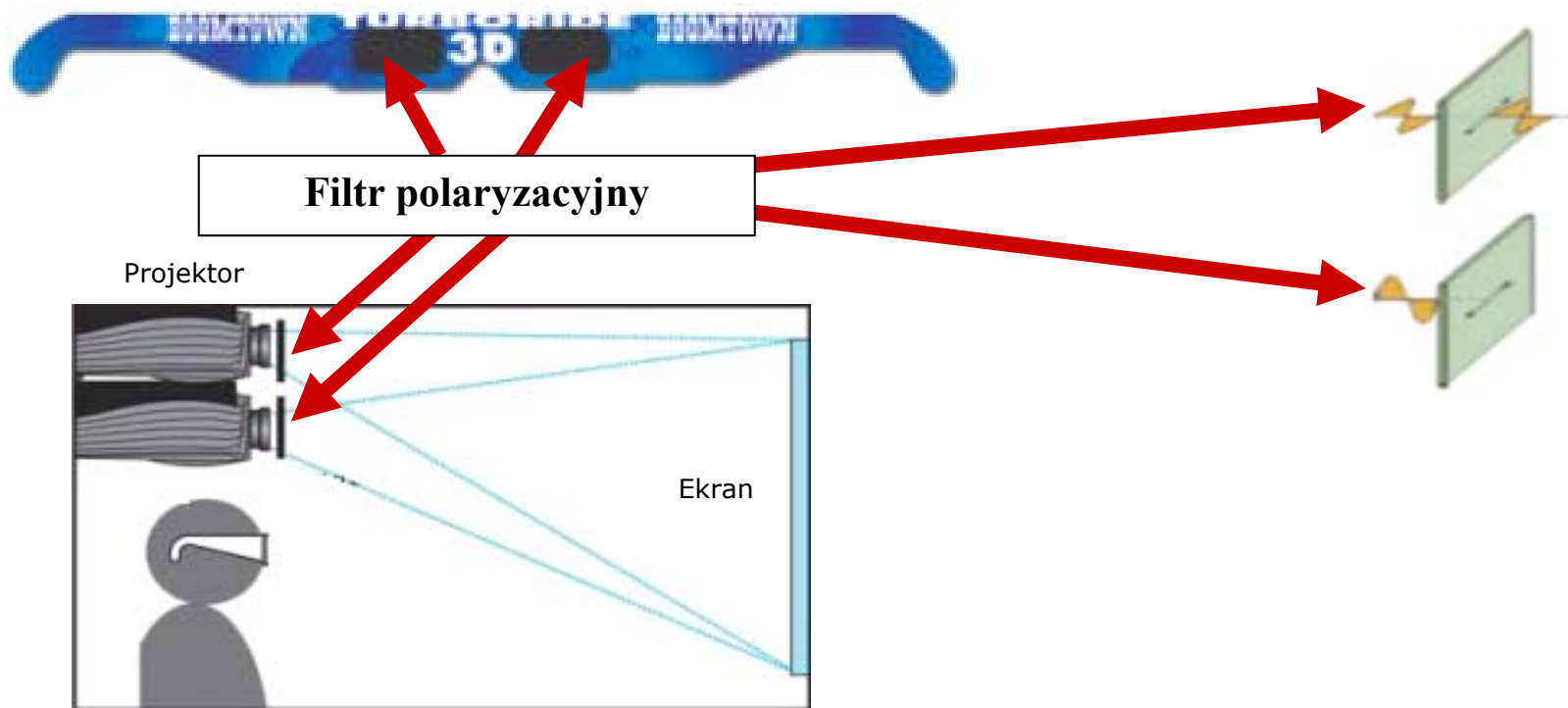
Okulary migawkowe



Wyświetlacze polaryzacyjne

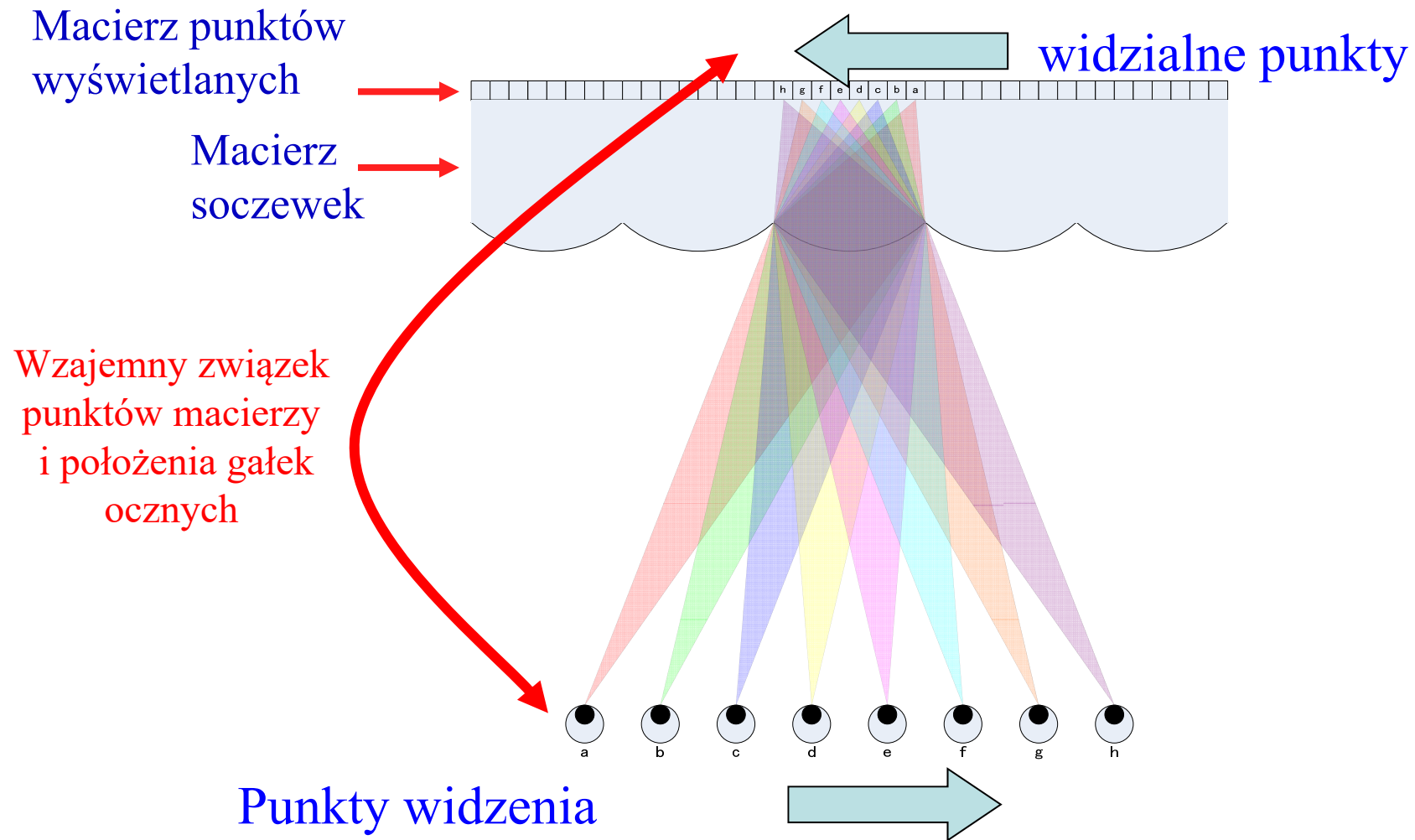


Projektory polaryzacyjne



Specjalne ekrany !!

Monitory autostereoskopowe dla wielu widzów bez okularów





Przykład

- Obiekt o wysokości $H=3$ metrów znajduje się na płaszczyźnie prostopadłej do osi optycznych pary kamer i odległej od środków optycznych tych kamer o $z=10$ metrów. Długość ogniskowa obu kamer wynosi $F=20$ mm. Przetworniki obu kamer składają się z 4000×2000 elementów (pikseli), okres próbkowania (pionowy i poziomy) wynosi $T_s=5$ μm . Baza stereoskopowa układu dwóch kamer wynosi $B=10$ cm.
A) Wyznaczyć rozbieżność d (liczoną w okresach próbkowania).
B) Jaką część (w procentach) wysokości przetwornika zajmuje obiekt?

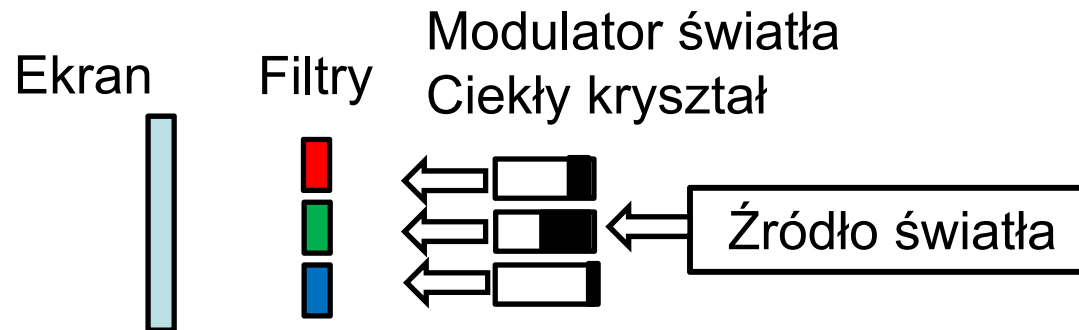
Zadanie

- Wyznaczyć maksymalną odległość od ekranu HD o wysokości 50 cm widza o zdrowym wzroku, taką , by z okularami polaryzacyjnymi rozróżniał 540 linii w obrazie.

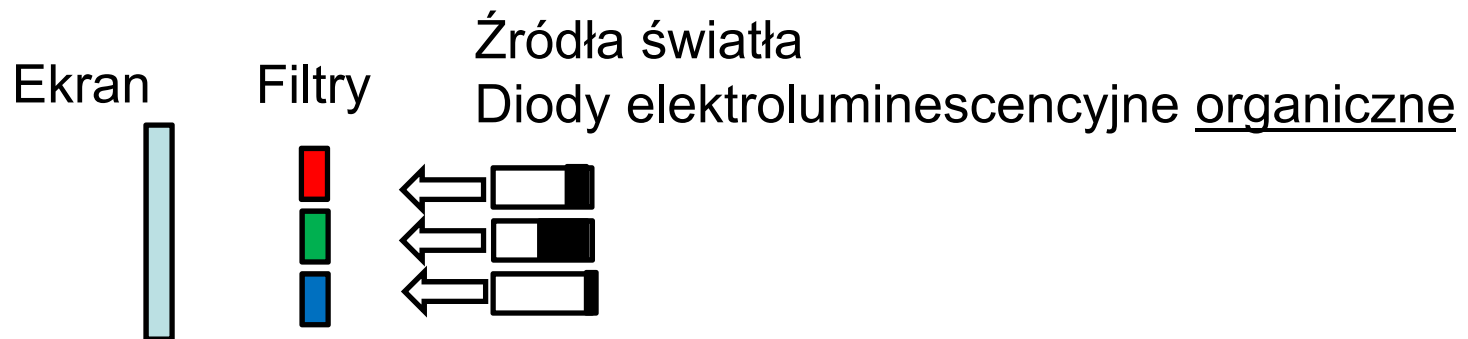
2.9. Monitory

Dwa podstawowe typy

LCD - Liquid-crystal display – Ekran ciekłokrystaliczny

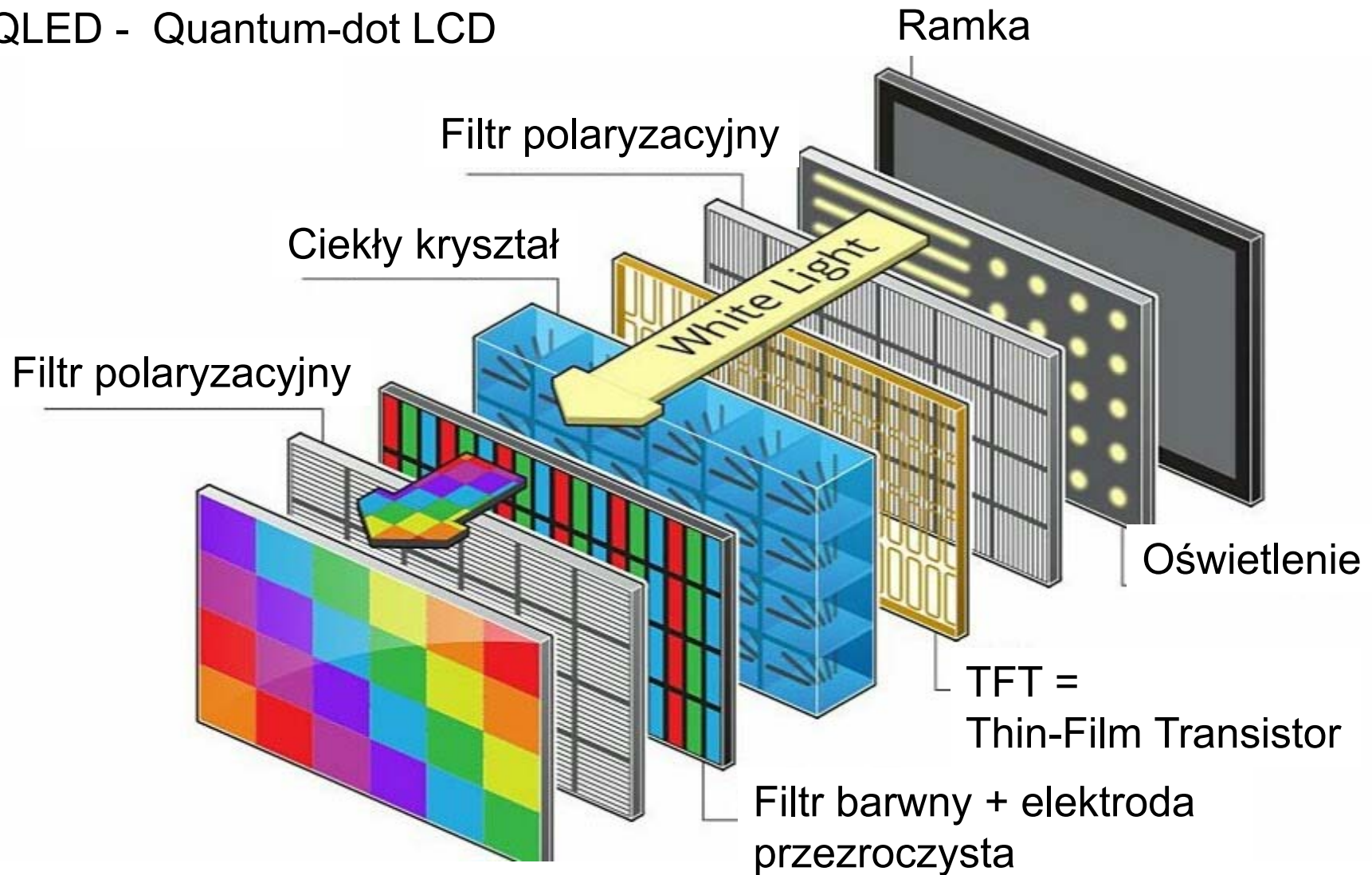


OLED – Organic light-emitting diode



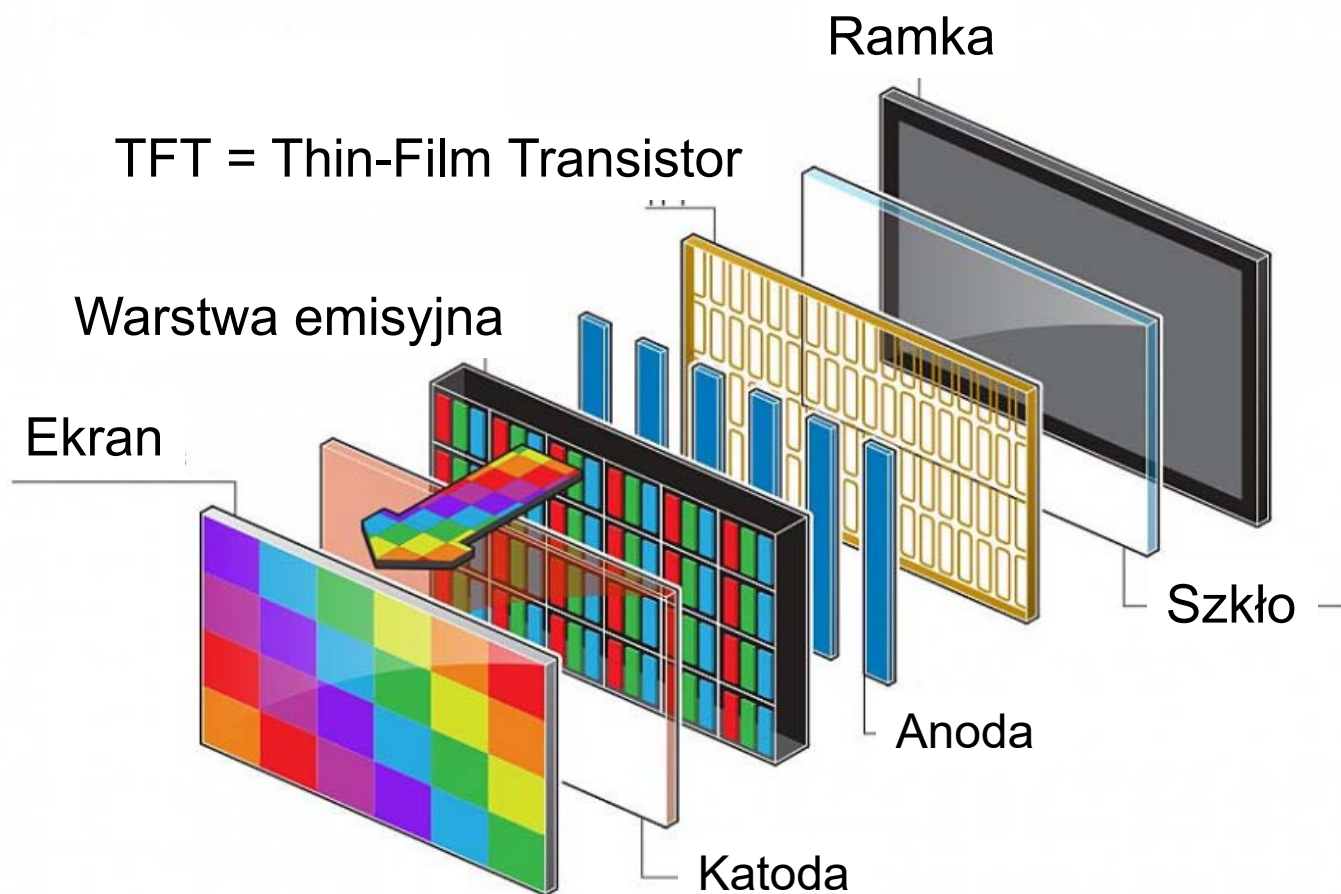
Monitory LCD

QLED - Quantum-dot LCD



Monitory OLED

AMOLED – Active-matrix OLED



Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego na studiach na kierunkach Elektronika i Telekomunikacja oraz Teleinformatyka.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

2.9. Ocena jakości obrazów

2.9.1. Metody subiektywne

Zalecenie ITU : BT.500

Średnia ocena grupy obserwatorów

MOS – *mean opinion score*

Po każdej prezentacji uczestnicy eksperymentów podają dla niej swoje oceny jakości w skali 5-stopniowej:

- 5 znakomity,
- 4 dobry,
- 3 dostateczny,
- 2 kiepski,
- 1 zły.

Alternatywnie, również w skali 5-stopniowej, oceniany jest poziom zakłóceń:

- 5 niezauważalne,
- 4 zauważalne, ale niedokuczliwe,
- 3 lekko dokuczliwe,
- 2 dokuczliwe,
- 1 bardzo dokuczliwe.

Metody subiektywne (c.d.)

metoda podwójnego wymuszenia i ciągłej skali jakości (**DSCQS** — *double stimulus continuous quality scale method*)

metoda pojedynczego wymuszenia i ciągłej skali jakości (**SSCQS** — *single stimulus continuous quality scale method*).

2.9.2. Metody obiektywne

$$PSNR[dB] = 10 \log \frac{N \cdot 255^2}{\sum_{i=1}^N e_i^2}$$

PSNR – peak signal-to-noise ratio

e_i - różnica (błąd) między odpowiednimi próbkami
w obrazie oryginalnym i badanym

N – liczba punktów w obrazie
(numerownych indeksem i)

Obraz oryginalny
PSNR = ∞ dB



**Do każdego punktu dodano 1
PSNR = 48,13 dB**



**Do każdego punktu dodano 2
PSNR = 42,11 dB**



Od $(1/16^2)$ punktów odjęto 32
PSNR = 42,11 dB



Przykład 2.8

- 10-bitowe próbki obrazu monochromatycznego reprezentują poziomy szarości i mają wartości z przedziału od 6 do 243. Do każdej próbki tego monochromatycznego obrazu oryginalnego dodano liczbę 5. Wyznacz wskaźnik *PSNR* dla tak przekształconego obrazu.

Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimedków

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimedków* prowadzonych przez prof. Marka Domańskiego.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

3. Wprowadzenie do przetwarzania obrazów

3.1. Pojęcie i zadania cyfrowego przetwarzania obrazów



Pojęcie cyfrowego przetwarzania obrazów

		wyjście	
		obraz	opis
wejście	opis	grafika komputerowa	
	obraz	cyfrowe przetwarzanie obrazów	analiza obrazu

3.2. Zadania przetwarzania obrazów

- Poprawa jakości (subiektywnej) obrazów (*image enhancement*)
- Rekonstrukcja obrazów (usuwanie zniekształceń i zakłóceń) (*image restoration*)
- Transformacje obrazów (zmiany reprezentacji) (*image transformations*)
- Kompresja (*compression*)
- Odtwarzanie obrazów z rzutów

Zadania przetwarzania obrazów (c.d.)

- Fuzja obrazów, sklejanie (stitching)
- Ekstrakcja cech (np. krawędzi)
- Segmentacja
- Śledzenie obiektów
- Analiza semantyczna, analiza zachowań
- Streszczanie wizji

- Grafika komputerowa
- Animacja komputerowa
- Tworzenie sztucznej rzeczywistości

3.3. Przykładowe zadania przetwarzania obrazów

Śledzenie obiektów

Object tracking

- Śledzenie obiektów jest procesem, w którym algorytm śledzi ruch obiektu w kolejnych ramkach wizji.
- Pojęcie dotyczyć może też ciągu nieruchomych obrazów (np. z różnych punktów).

Śledzenie obiektów

Object tracking

Podstawowe etapy:

- Wykrycie obiektu – algorytm wykrywa i klasyfikuje obiekt – tworzy prostokątne obramowanie (bounding box) obiektu.
- Każdy obiekt otrzymuje unikalną etykietę.
- Śledzenie wykrytego obiektu w kolejnych ramach wizji i zapamiętanie odpowiedniej informacji

Śledzenie obiektów

Object tracking

Podstawowe etapy:

- Wykrycie obiektu – algorytm wykrywa i klasyfikuje obiekt – tworzy prostokątne obramowanie (bounding box) obiektu.
- Każdy obiekt otrzymuje unikalną etykietę.
- Śledzenie wykrytego obiektu w kolejnych ramach wizji i zapamiętanie odpowiedniej informacji

Śledzenie obiektów - typy

Dwa podstawowe typy:

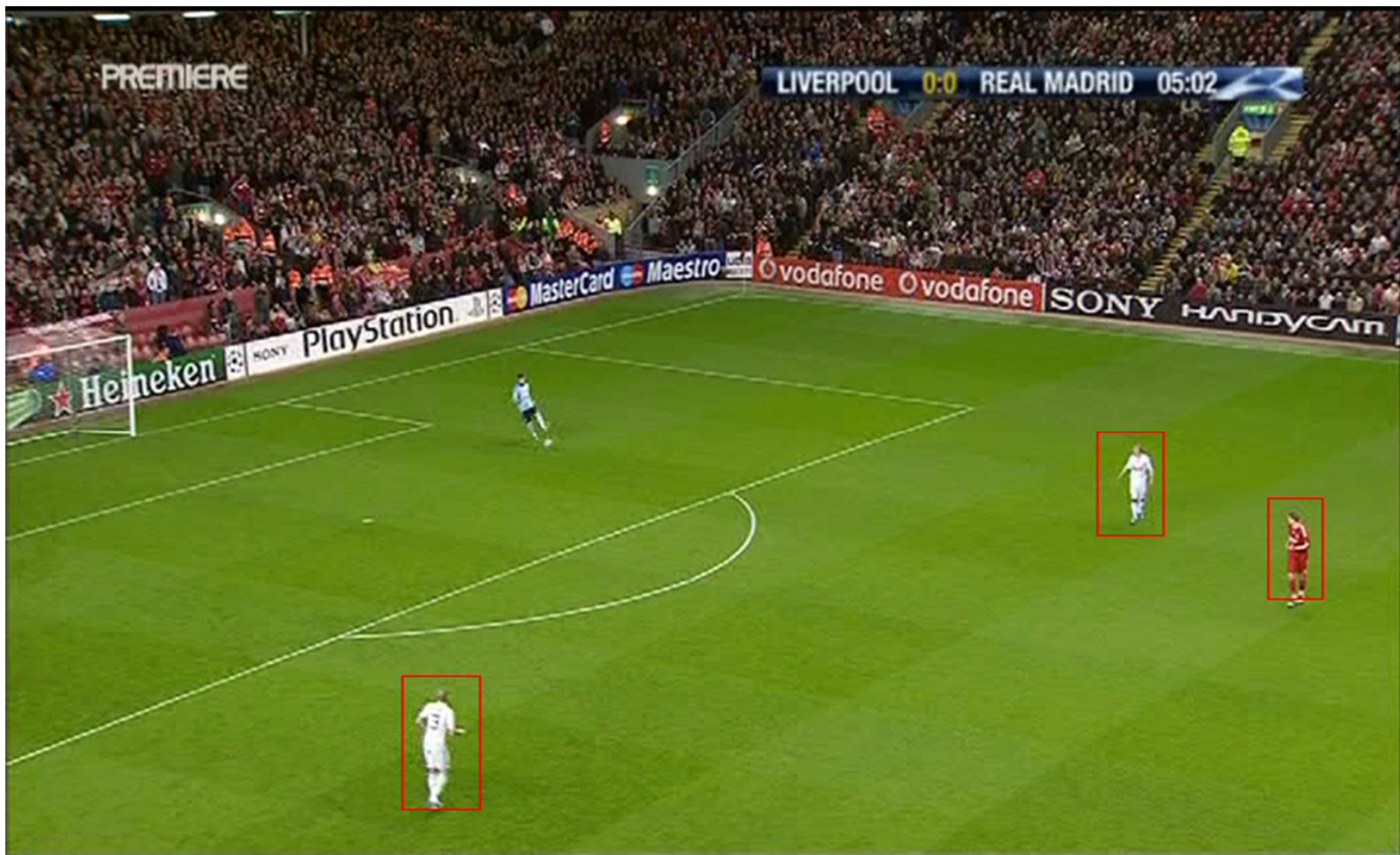
- Śledzenie w obrazach – *Image tracking* – automatycznie rozpoznaje i śledzi obiekty.
Przykład: Trzymamy w ręku kamerę, nasza ręka się porusza, ale algorytm utrzymuje prostokątne obramowanie wokół obiektu.
- Śledzenie wizji – *Video tracking* – automatycznie rozpoznaje i śledzi poszczególne obiekty w sekwencji wizyjnej.

Śledzenie obiektów - problemy

- Przesłanianie – *occlusion*
- Zaśmiecanie tłem – *background clutter*

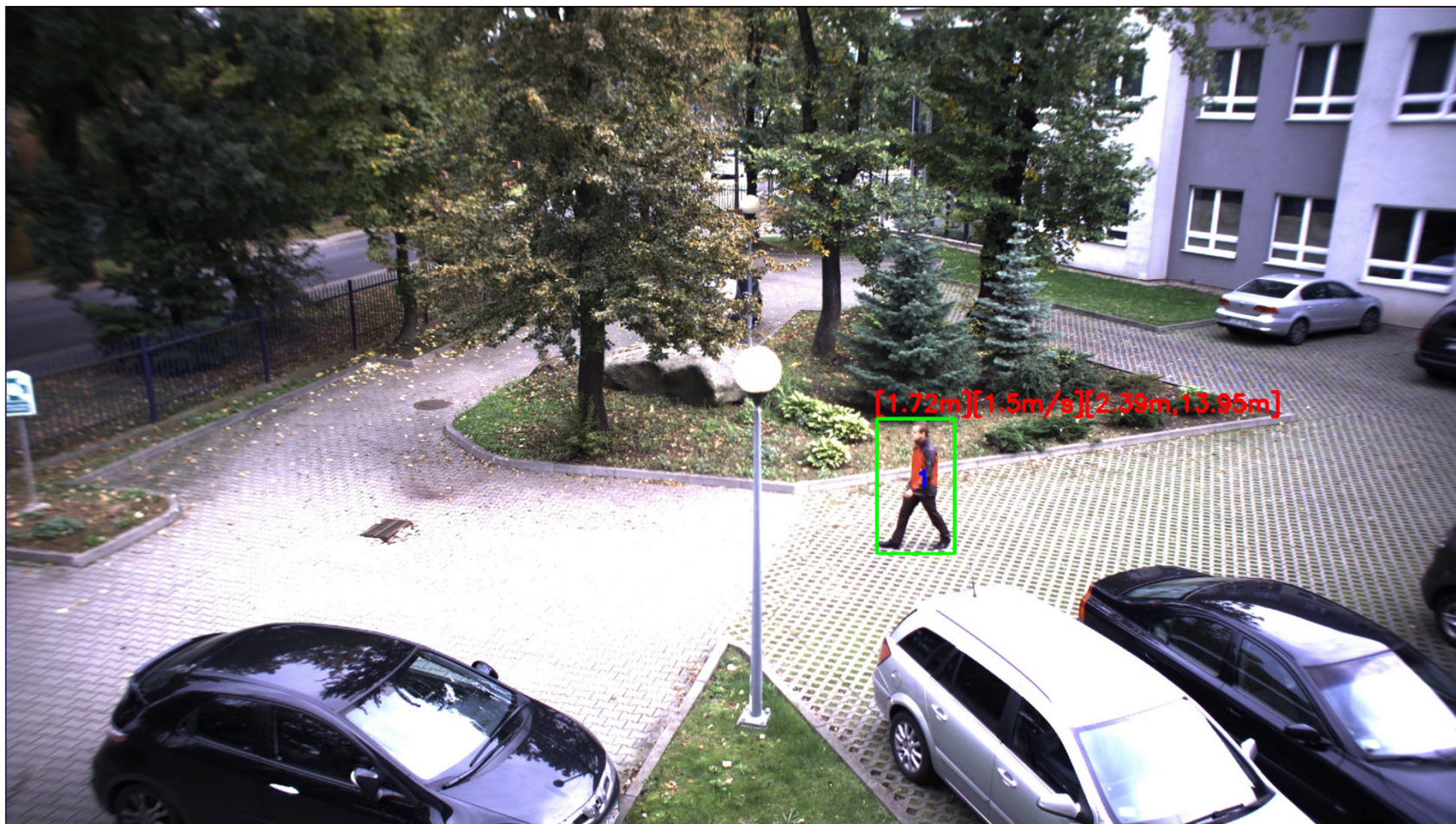
Śledzenie

Video tracking



Śledzenie

Video tracking



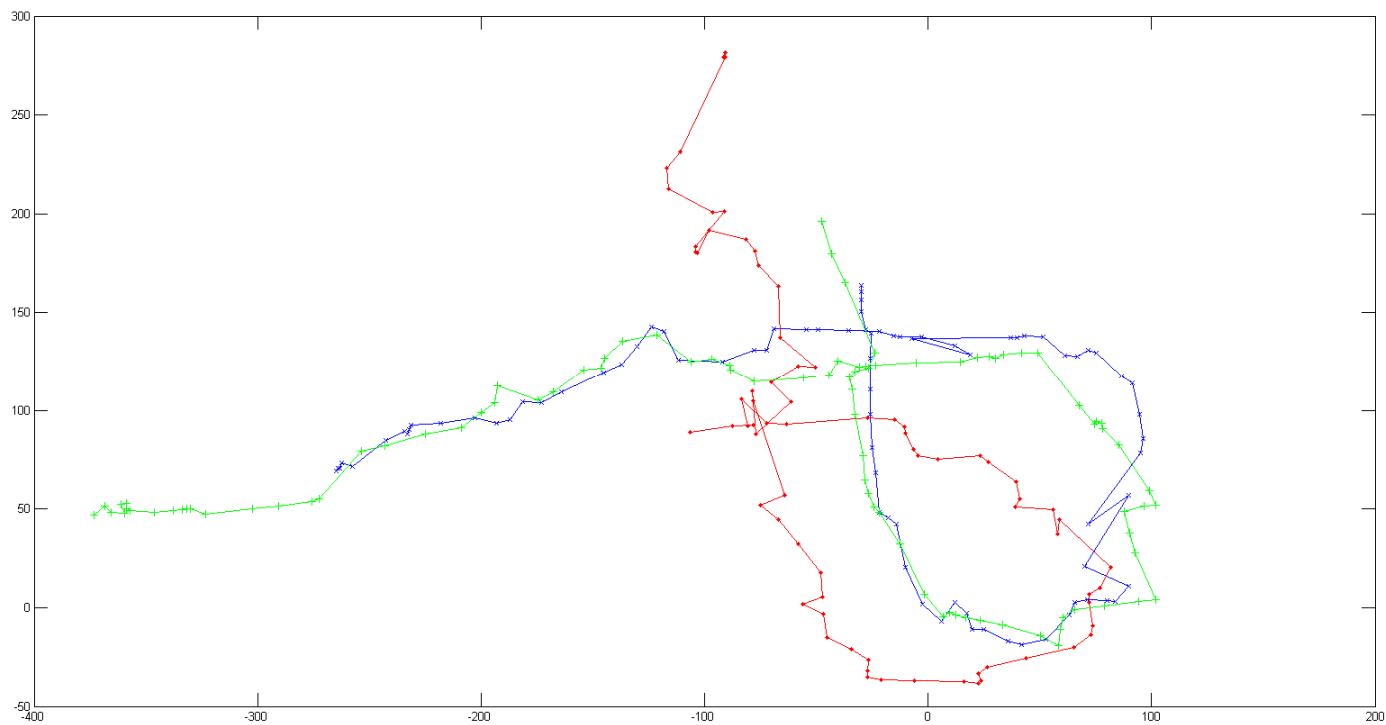
Trajektoria ruchu

Motion trajectory

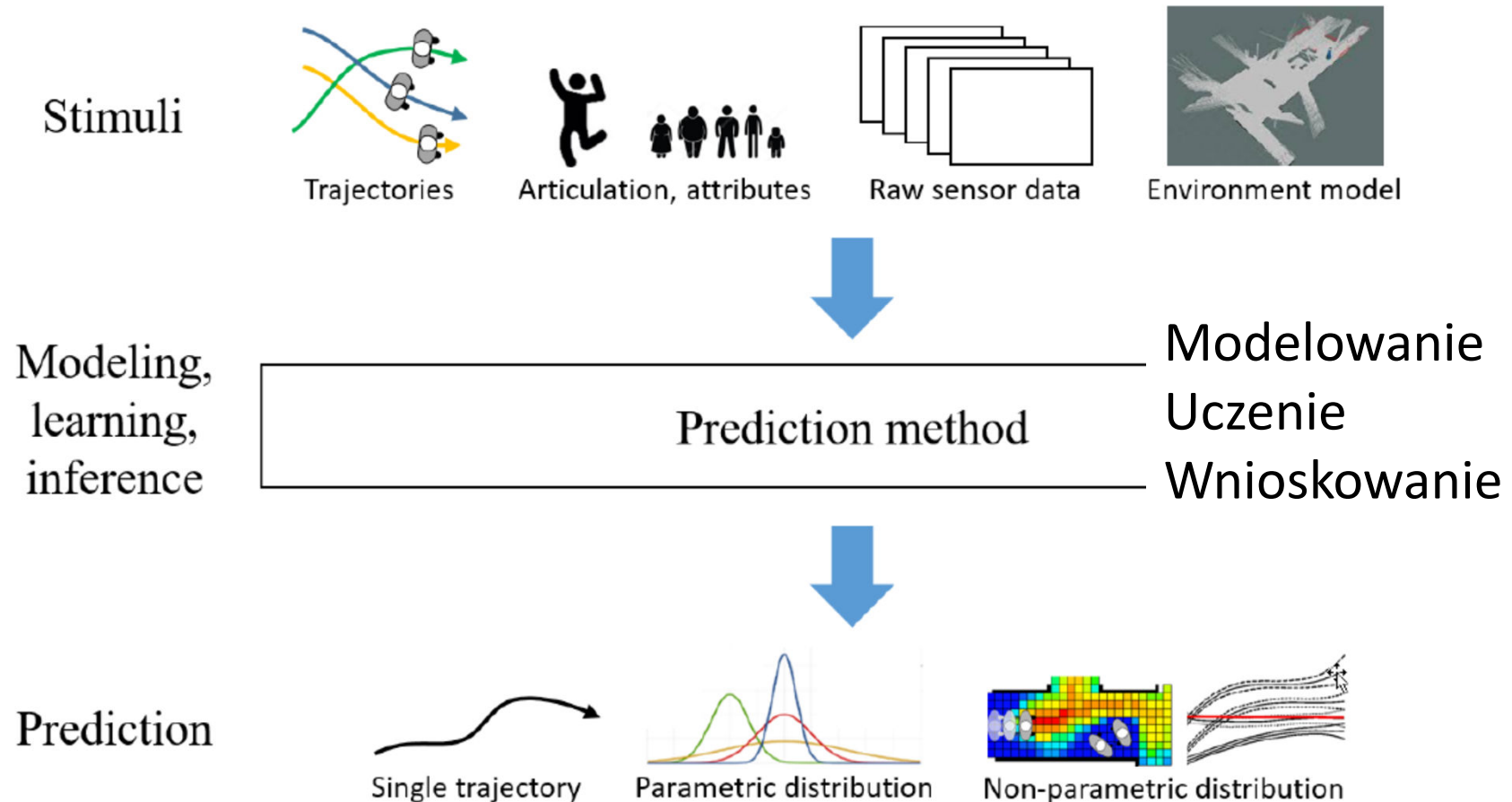
- Rejestracja kolejnych położeń charakterystycznego punktu w obrazie obiektu (np. jego środka określonego w pewien sposób).
- Predykcja trajektorii – przewidywanie dalszej trajektorii ruchu na podstawie dotychczasowego ruchu.

Trajektoria ruchu

Przykład



Predykcja trajektorii ruchu

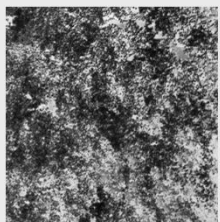
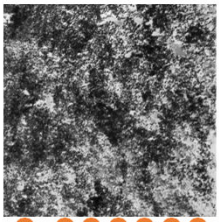
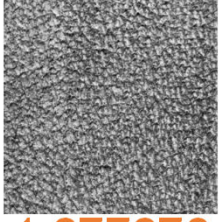
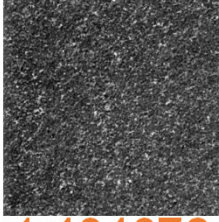
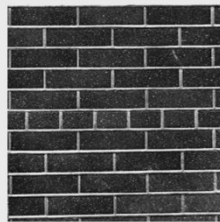
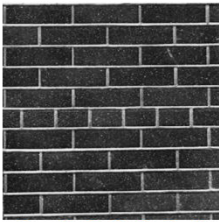
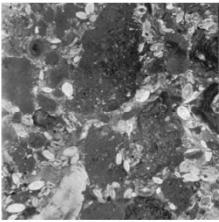
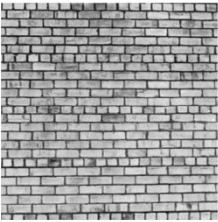
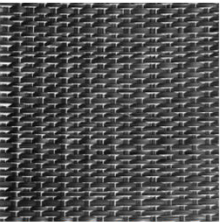
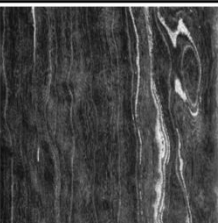
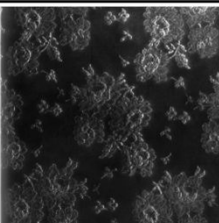
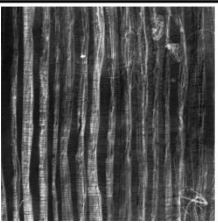
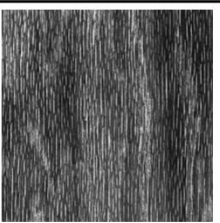


Analiza i klasyfikacja tekstur

- W przetwarzaniu obrazów:

Teksturę określa się jako funkcję
przestrzennego rozkładu barw punktów.

Analiza i klasyfikacja tekstur

 e	 0.000000	 1.012509	 1.277278	 1.434672	 1.440774	 1.595892
 e	 0.000000	 1.849781	 2.168069	 2.273702	 2.309718	 2.364708
 e	 0.000000	 1.298014	 1.958568	 1.959547	 2.083975	 2.147551

Analiza wysokiego poziomu

- Na podstawie analizy cech dokonuje się **analizy semantycznej**:
 - Klasyfikacja obiektów: np. człowiek, pies, drzewo,
 - Klasyfikacja zachowań osób (marsz, bieg, bójka, omdlenie, ...).
 - Rozpoznawanie obiektów (np. konkretnych osób, rozpoznawanie twarzy, odczyt tablic rejestracyjnych pojazdów)

Analiza zachowań

EVENT DETECTION



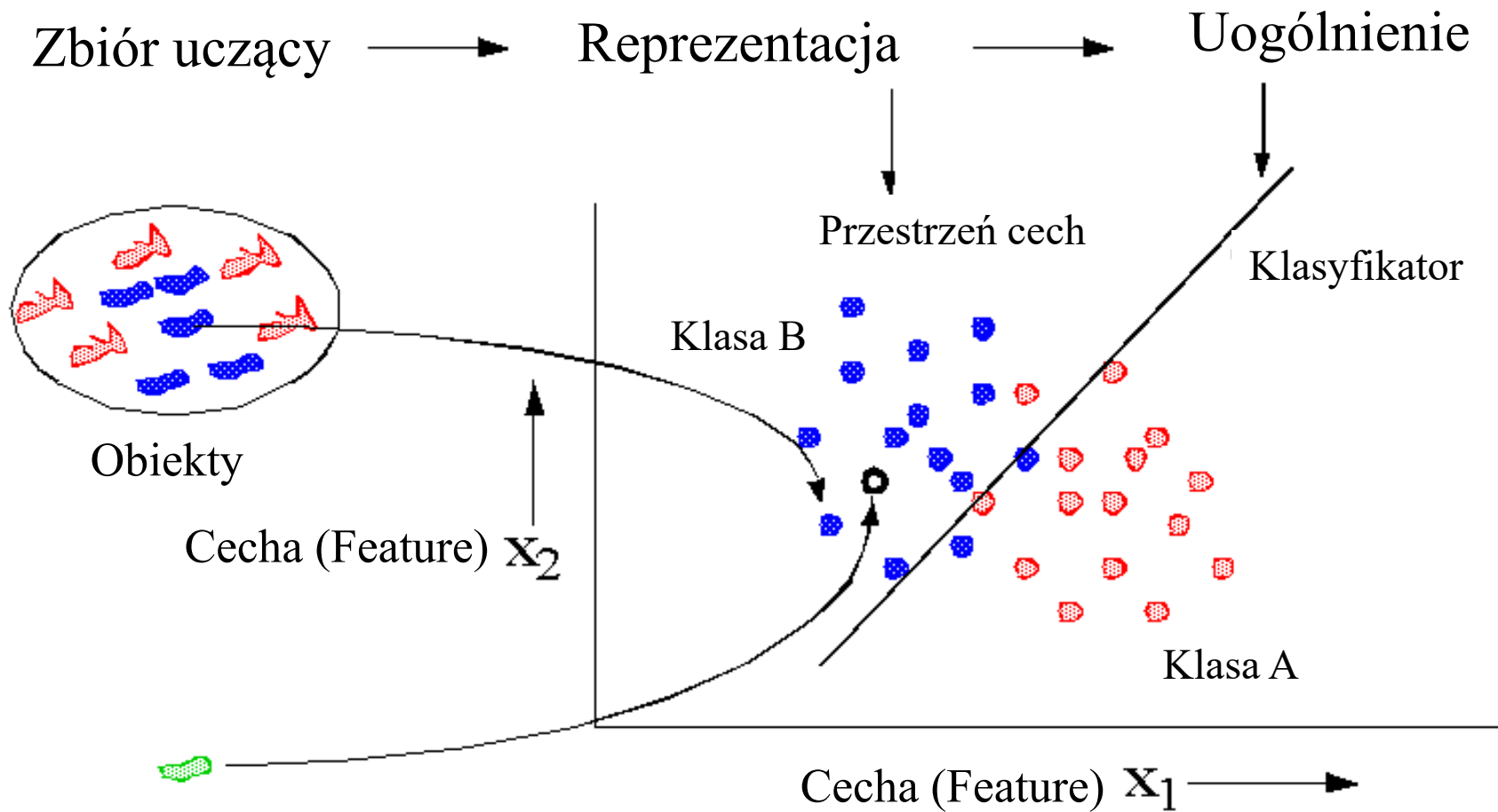
Rozpoznawanie

Recognition

- Stosuje się metody zgłębiania wiedzy (*data mining*)
- Uczenie:
 - wyznaczenie cech dyskryminatywnych
 - wyznaczenie reguł klasyfikacji
- Wyznaczenie cech dla wszystkich obiektów
- Klasyfikacja obiektów

Rozpoznawanie

Recognition



Cechy: np. stosunek długości do szerokości (x_1), chroma CR (x_2)

Rozpoznawanie

Recognition

- W ostatnich latach wykorzystanie metod sztucznej inteligencji (*artificial intelligence*)
- W tym głębokich sieci neuronowych (*deep neural networks*)

Analiza zachowań

EVENT DETECTION



Deskryptory

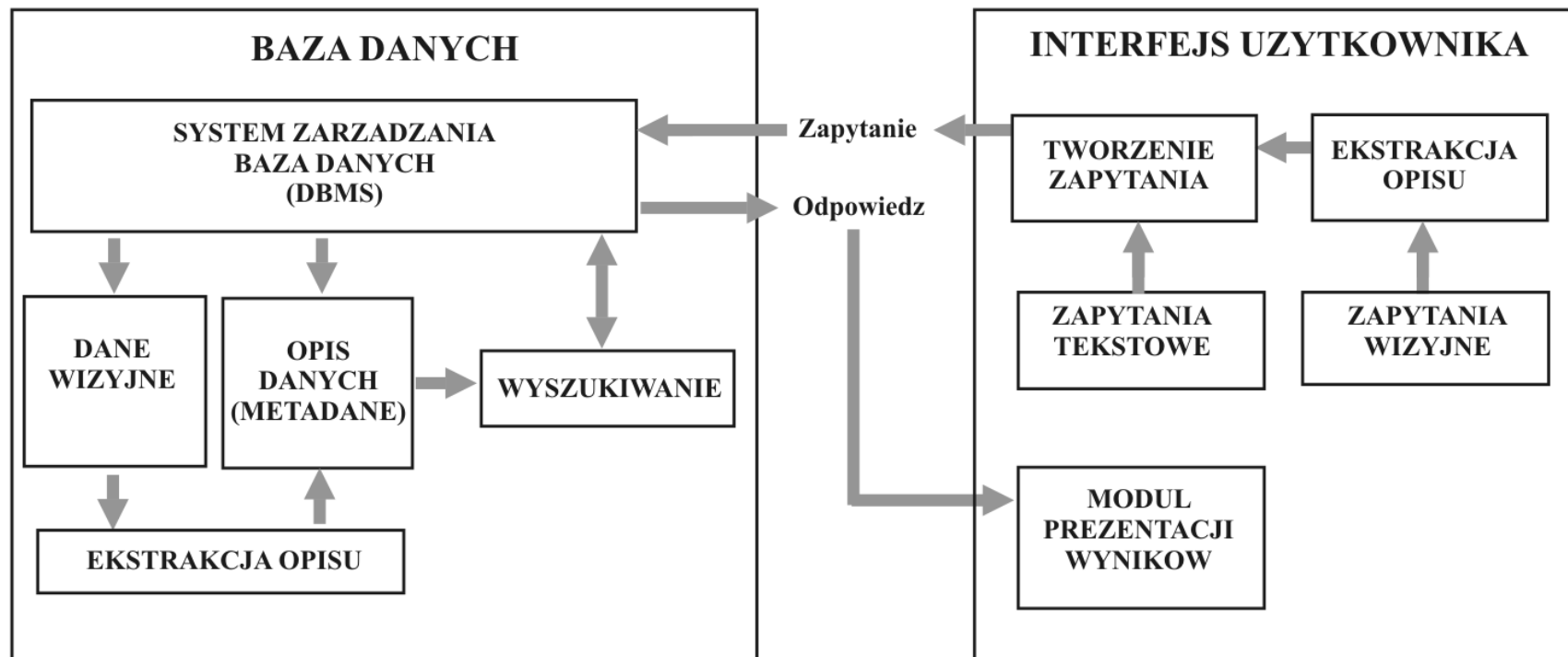
- Deskryptor to wektor liczb reprezentujących parametry poszczególnych cech)
- Istnieją normy dla deskryptorów – znormalizowane sposoby opisu cech obiektów, fragmentów obrazów, obrazów np. MPEG-7

Wykorzystanie deskryptorów

- Baza danych – indeksowanie
wyznaczenie wartości deskryptorów dla
poszczególnych obiektów x_i
 $x_i \leftrightarrow \underline{E}_i = [F_{1i}, F_{2i}, \dots, F_{Ni}]$
- Zapytanie obiektem y
 $y \leftrightarrow \underline{E}_y = [F_{1y}, F_{2y}, \dots, F_{Ny}]$
- Wyszukanie w bazie obiektu $x^* = x_i$, dla
którego
 $\| \underline{E}_i - \underline{E}_y \|$ jest minimalna

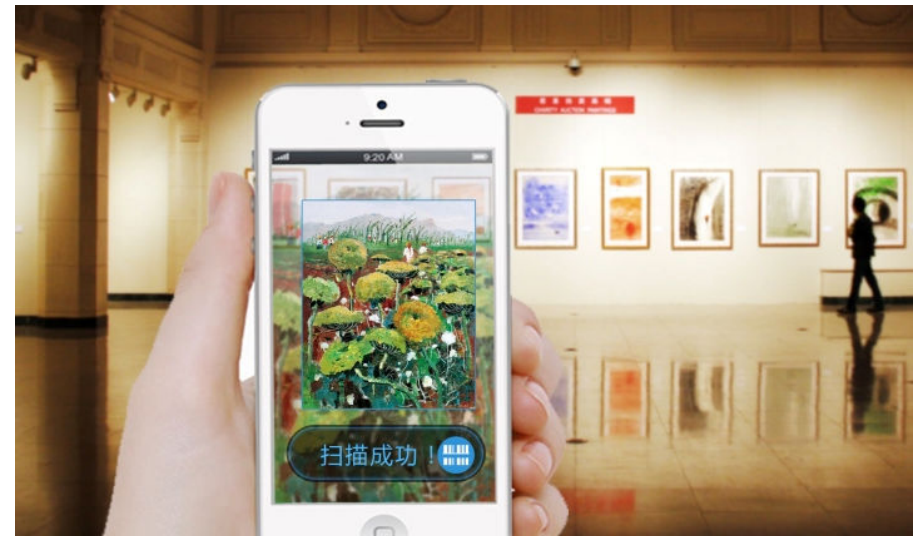
Wyszukiwanie danych multimedialnych

- Schemat systemu wyszukiwania na podstawie opisu zawartości danych (metadanych)



Compact descriptors for visual search (CDVS)

**Mobile visual search:
search for information about landmarks and arts**



Compact descriptors for visual search (CDVS)

Mobile visual search:

Applications in augmented reality

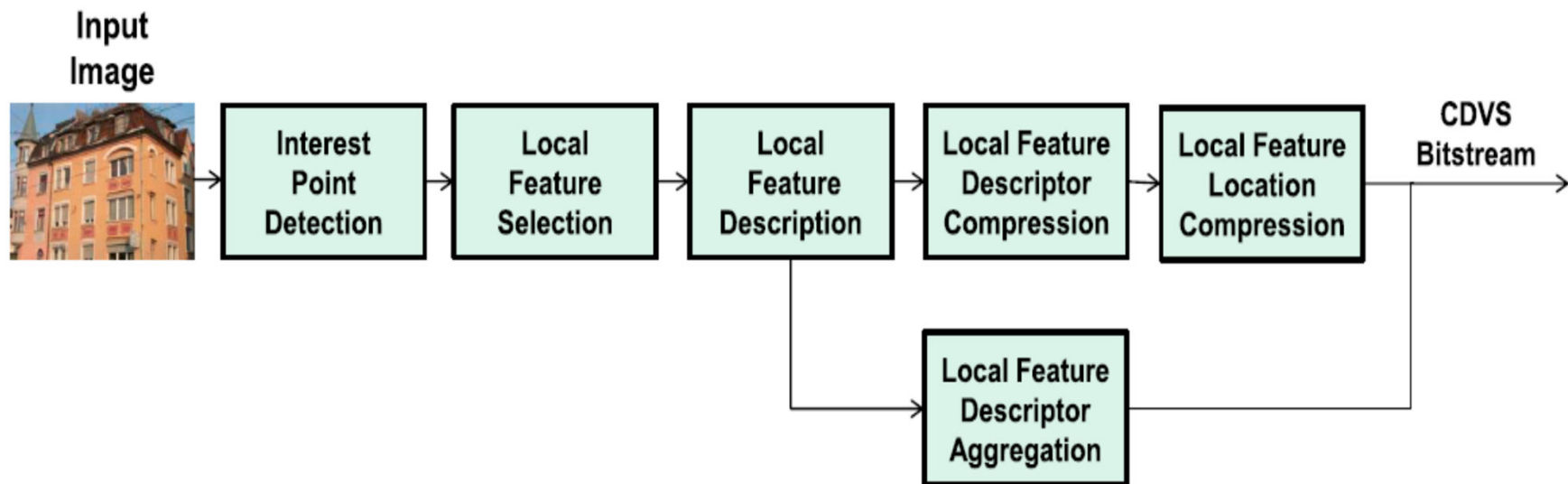


Compact descriptors for visual search (CDVS)

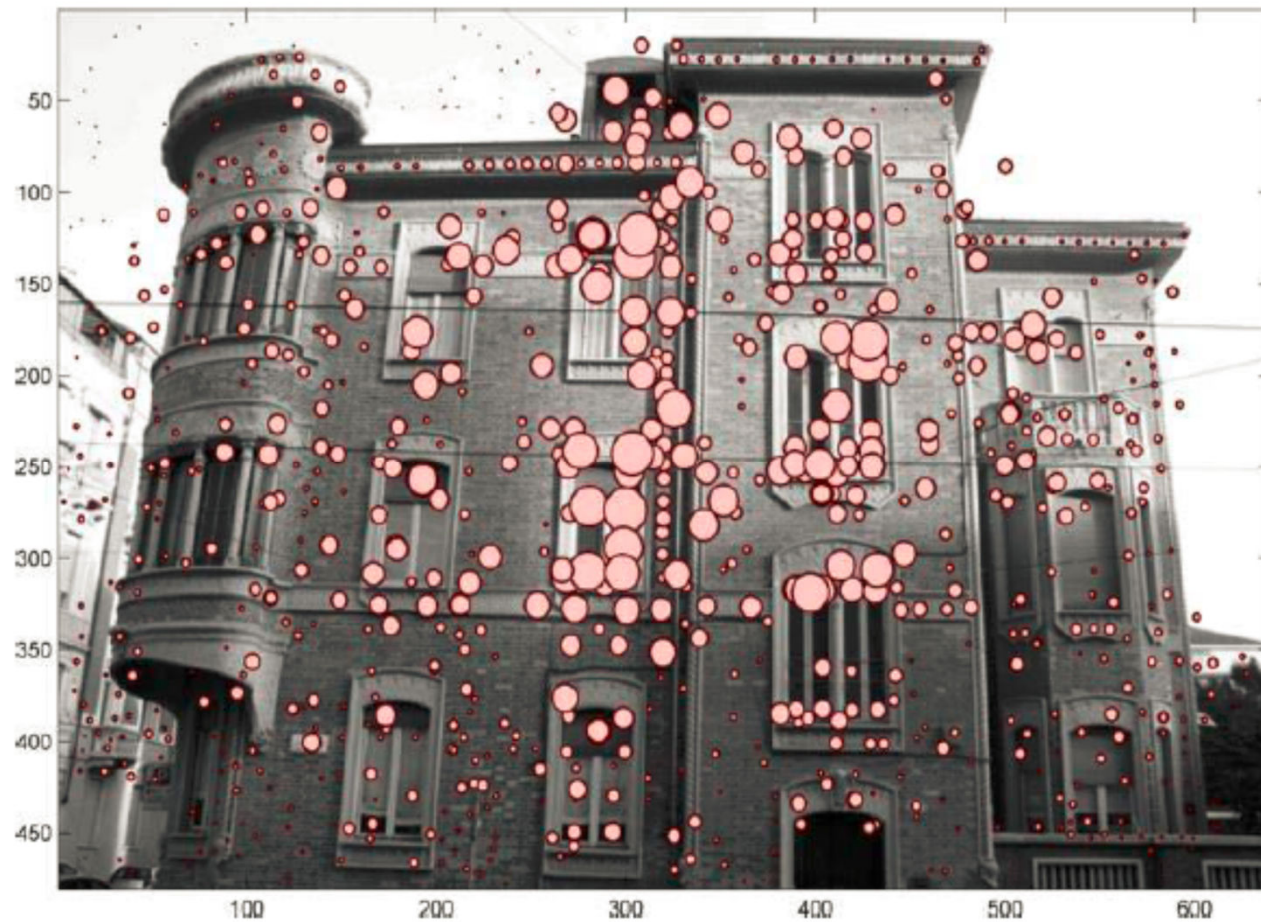
TV/IPTV – related applications



Compact descriptors for visual search (CDVS)



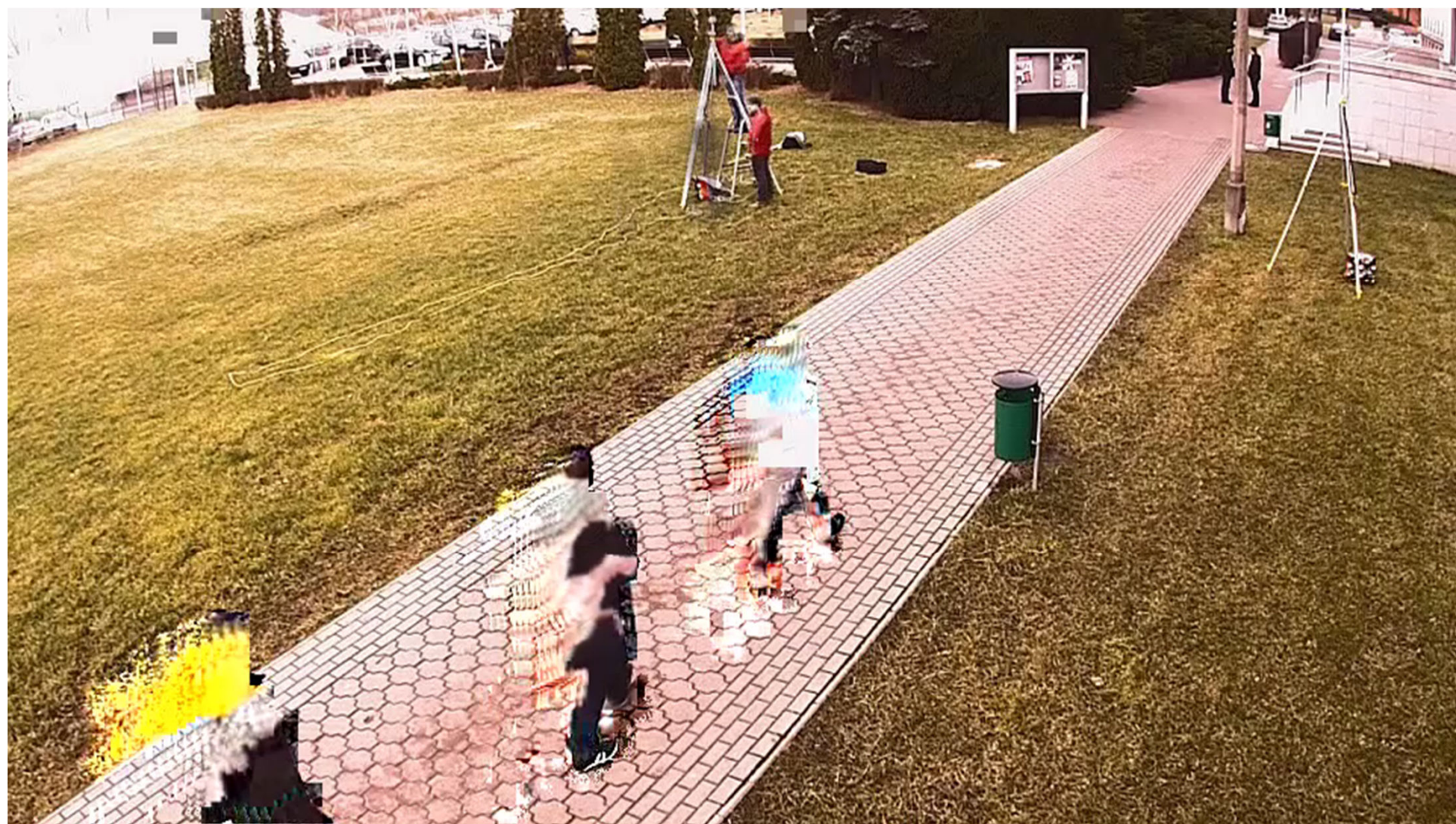
Compact descriptors for visual search (CDVS)

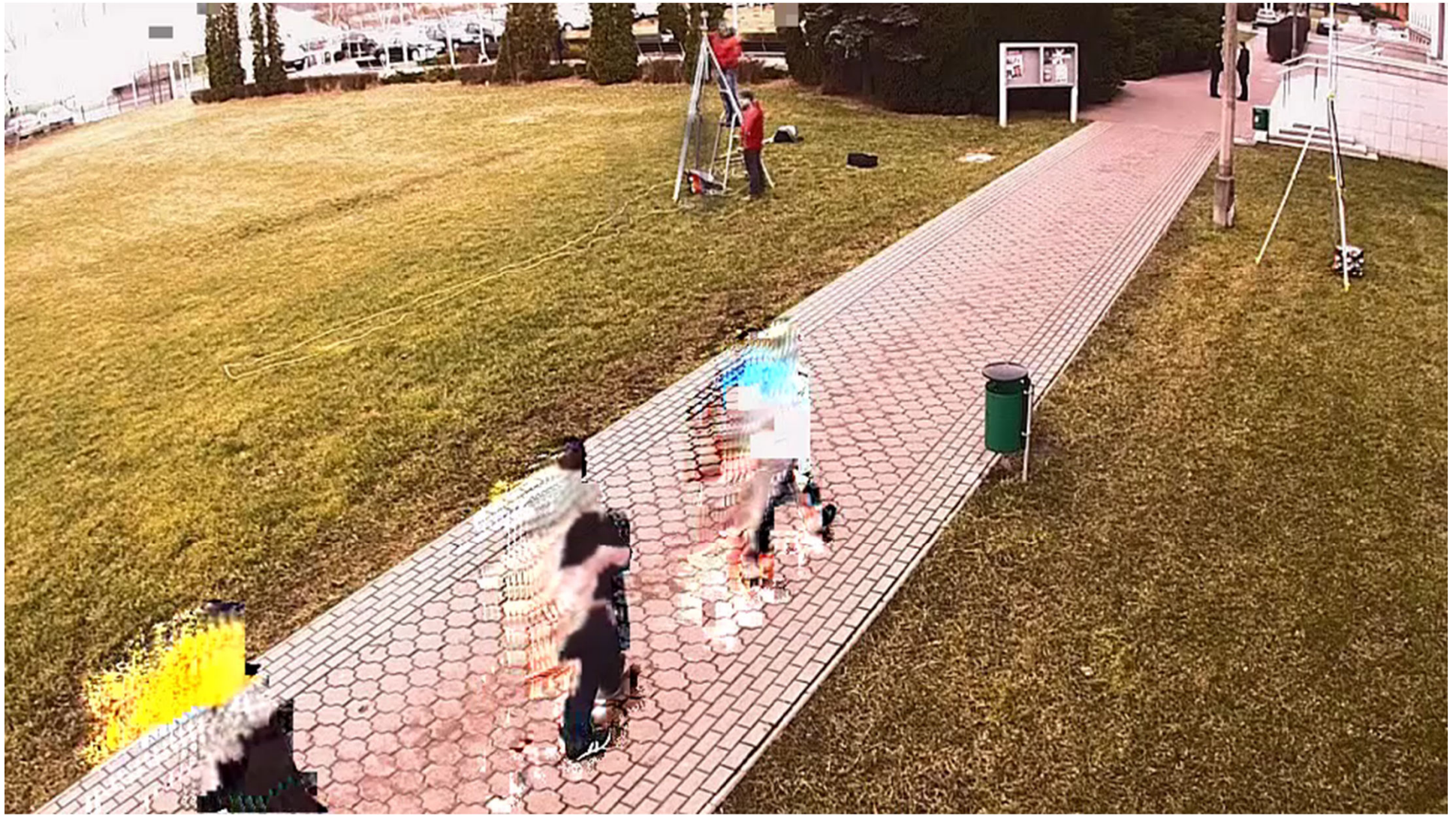


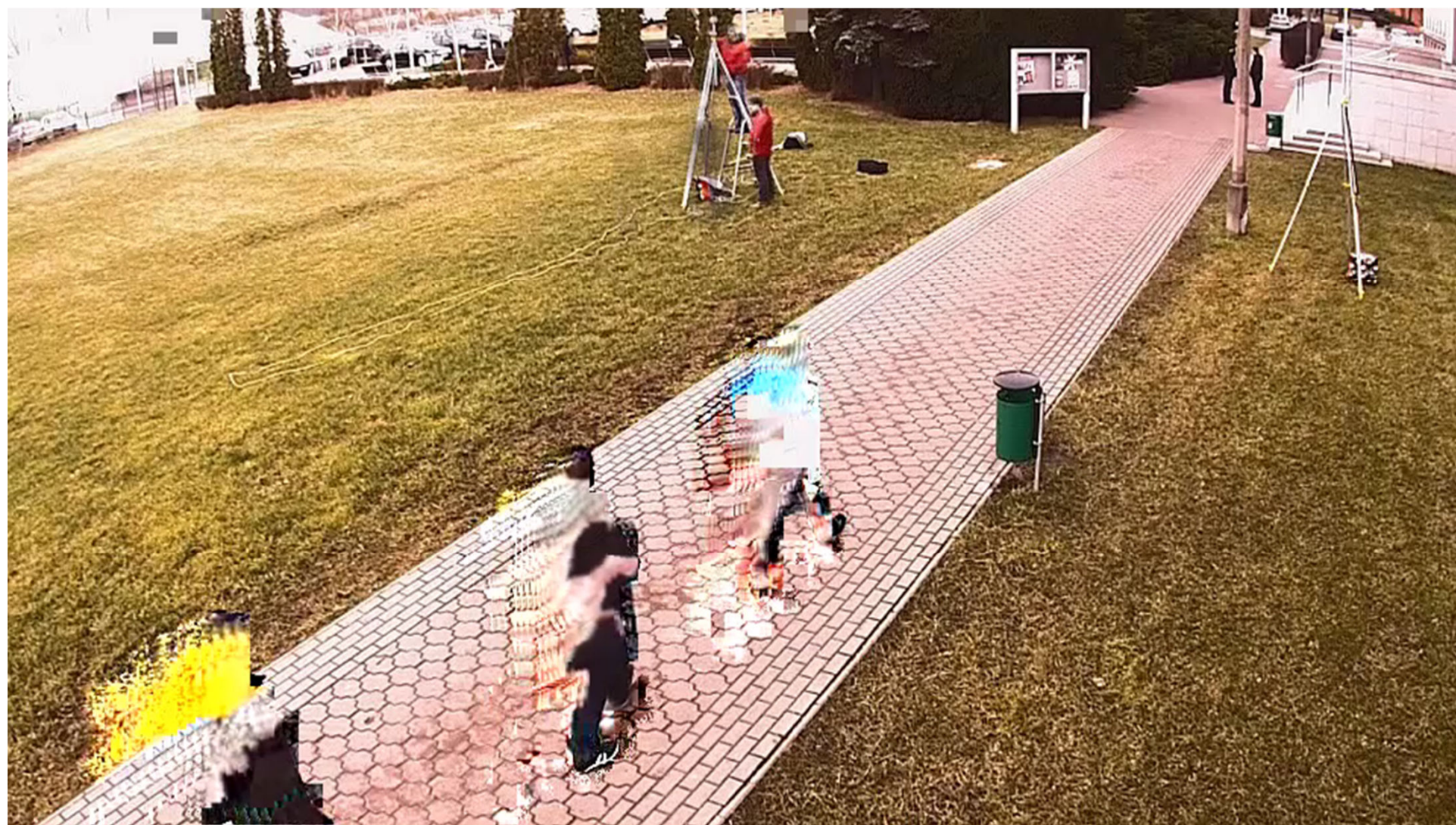
Streszczenie wizji

Video synopsis









Inne

- Rozproszone systemu dozoru wizyjnego
- Identyfikacja kamery
- Znakowanie wodne

Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów **Politechnika Poznańska**

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

3.4. Podstawowe operacje przetwarzania obrazów

Operacje mające na celu

poprawę subiektywnej jakości obrazu

i/lub usunięcie zakłóceń

i/lub dokonanie zmian w obrazach

3.4.1. Operacje punktowe

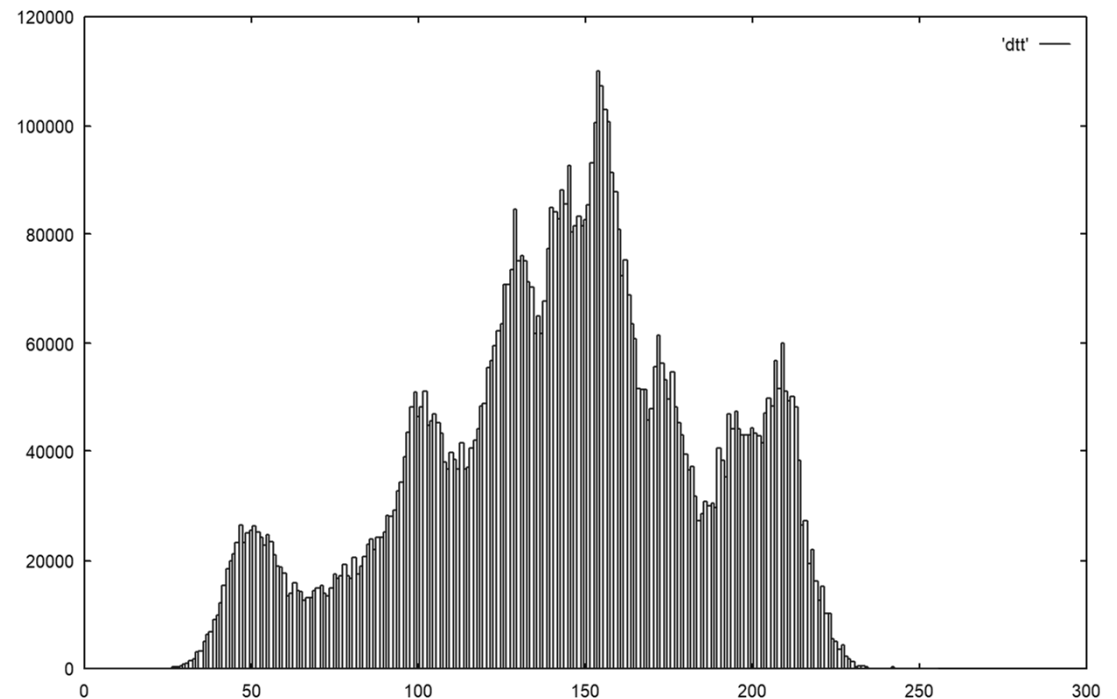
Grupa bardzo prostych operacji na obrazach

Histogramy

Histogram obrazu monochromatycznego:

$H(u)$ – liczba punktów obrazu o wartości u ,

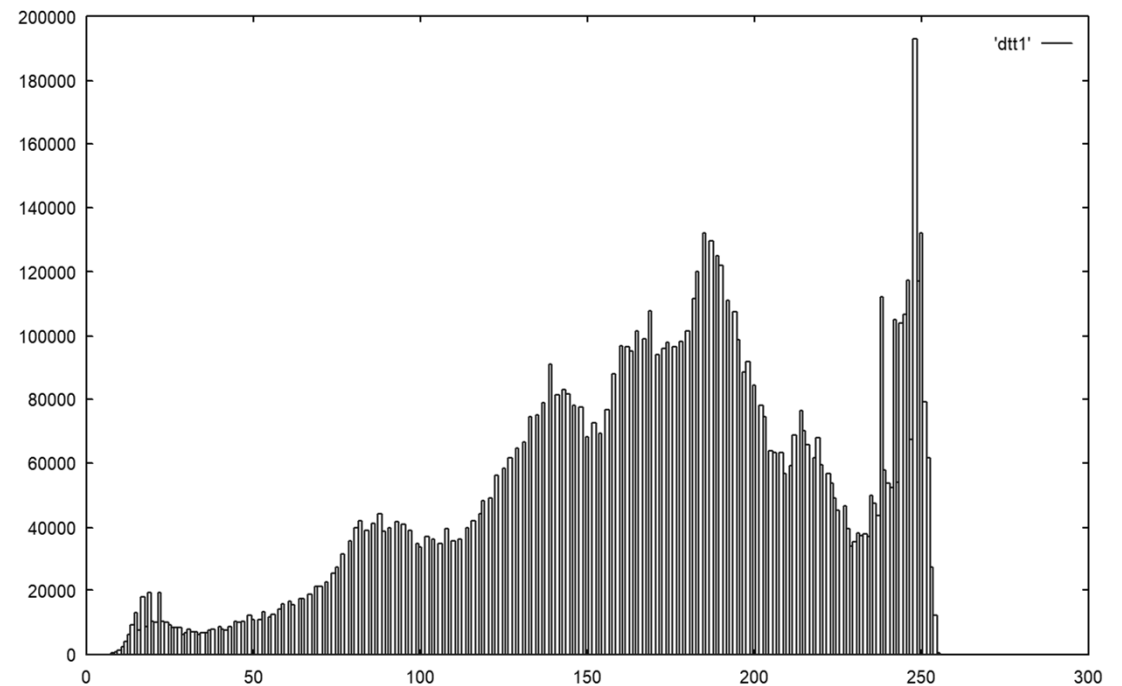
Dla 8-bitowej reprezentacji próbek: $0 \leq u \leq 255$.



Obrazowi o subiektywnie „dobrym kontraście” odpowiada histogram o znacznych wartościach $H(u)$ w całym przedziale (0,255)

Obraz zbyt kontrastowy – utrata szczegółów w najjaśniejszych partiach obrazu.

Histogram: „ucięty” dla wartości ok. 255.

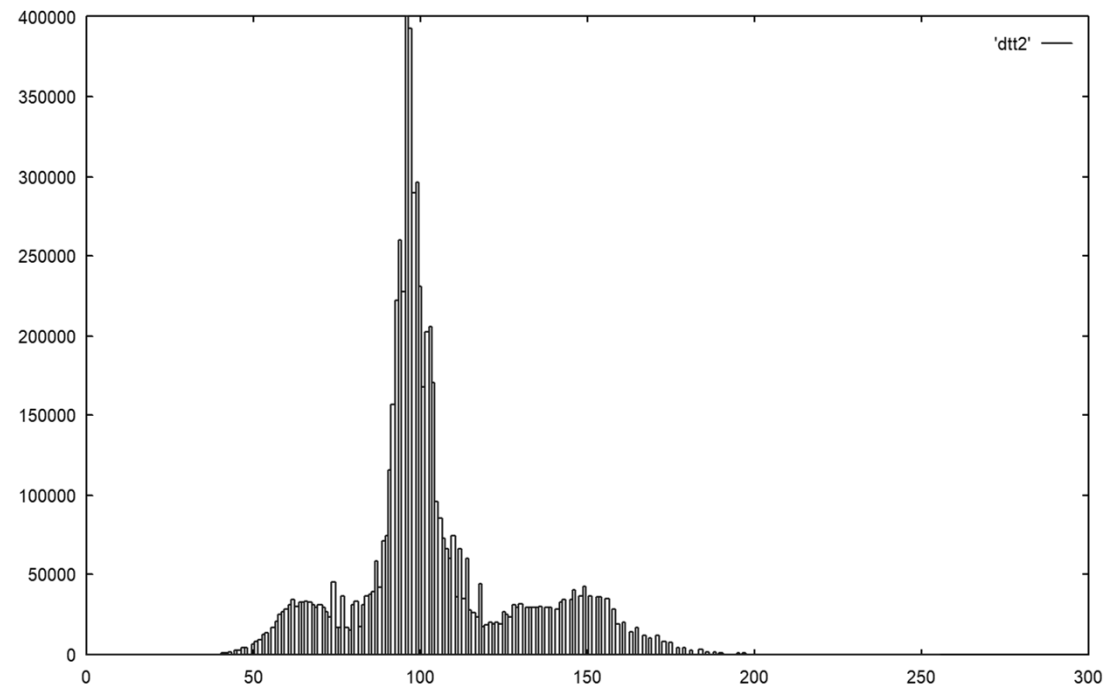


Obraz mało kontrastowy –

– trudno rozróżniać szczegóły, niekorzystne wrażenie estetyczne.

Histogram:

znaczące wartości histogramu występują tylko w części przedziału (0,255).



Analiza histogramu umożliwia automatyczną ocenę subiektywnej jakości kontrastu obrazu.

Operacje punktowe - definicja

Operacja punktowa $F[\cdot]$ przekształca obraz wejściowy $u(n_1, n_2)$ w obraz wyjściowy $y(n_1, n_2)$ zgodnie ze wzorem

$$y(n_1, n_2) = F [u(n_1, n_2)] .$$

Operacje punktowe mogą być interpretowane jako transformacje histogramów.

Obraz oryginalny

Drewniana
pagoda Horyuji
w Japonii.
Pochodzi z 7 w.



Dodanie liczby 70
do każdej próbki obrazu
(widoczne efekty
przekroczenia zakresu)



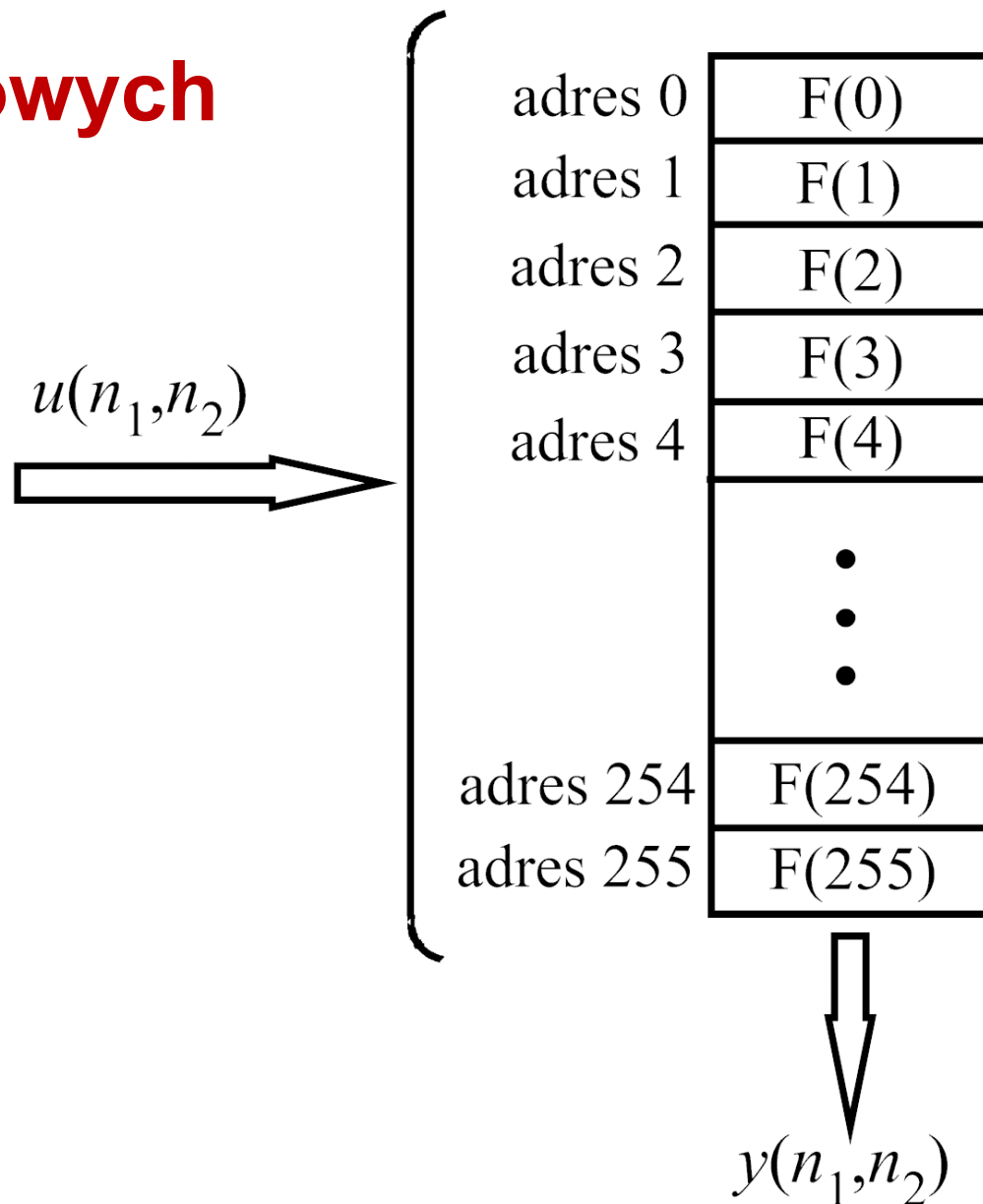
Dodanie liczby 70
do każdej próbki obrazu
z ograniczeniem
wyniku do 255



Inwersja obrazu
 $y = 255 - u$

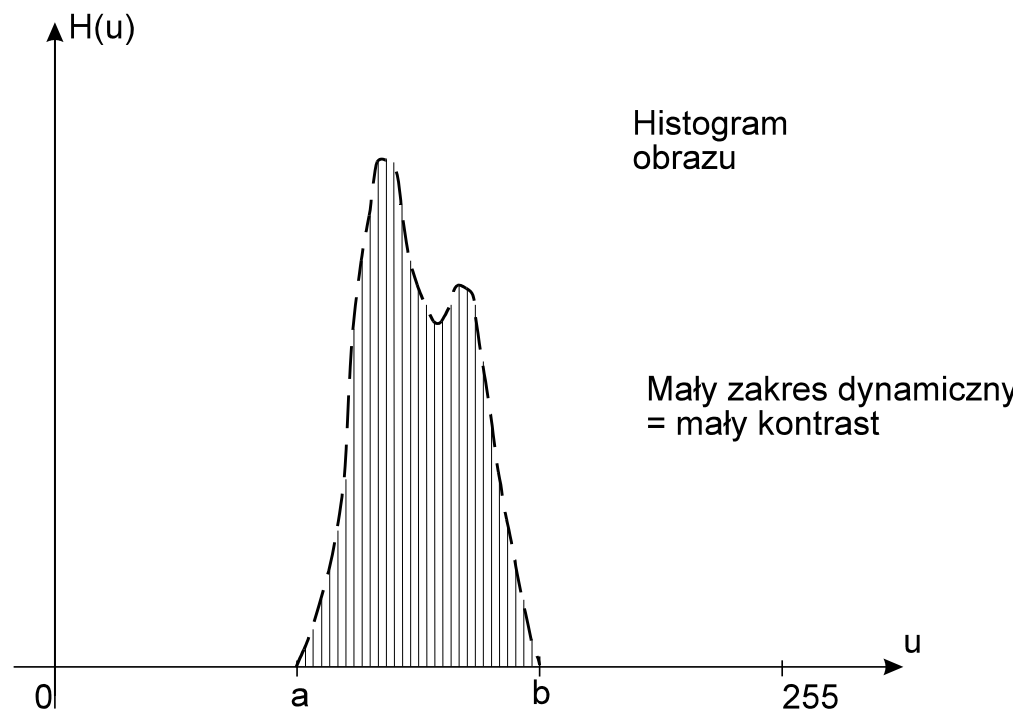


Implementacja operacji punktowych



Poprawa kontrastu

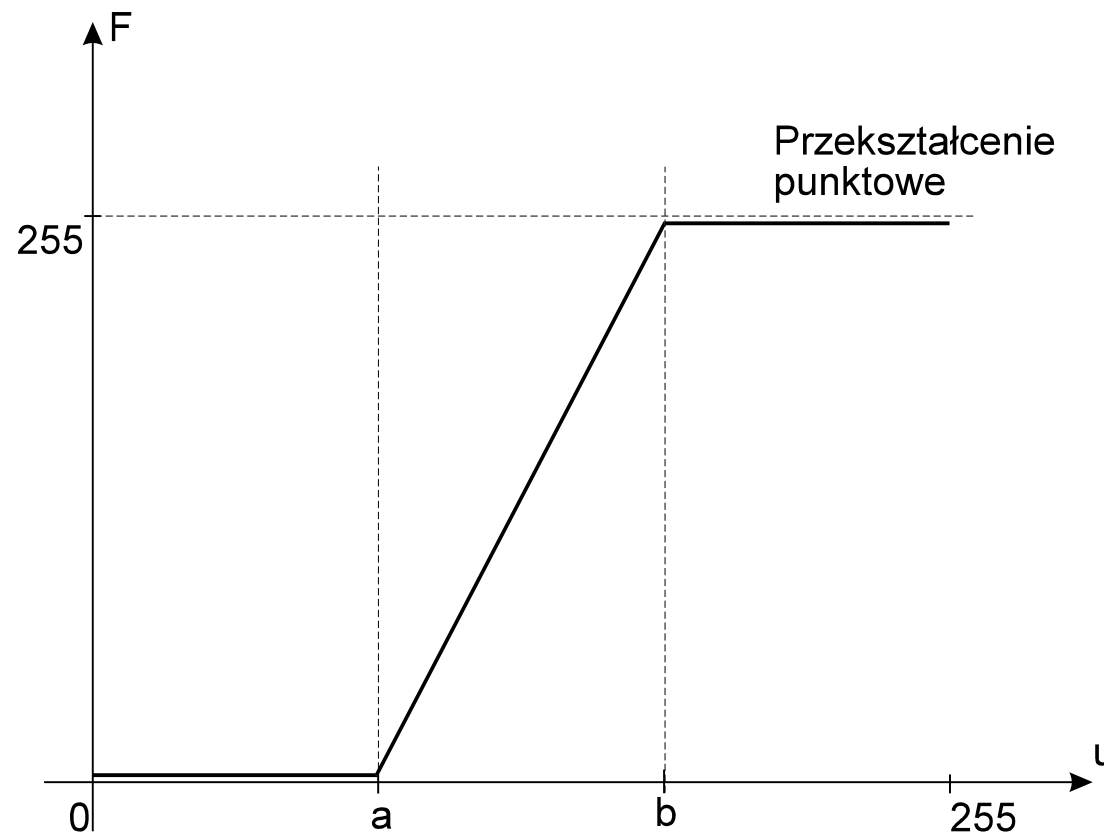
Jako grupa operacji punktowych.



Za pomocą odpowiedniej operacji punktowej poprawimy **subiektywną** jakość tego obrazu.

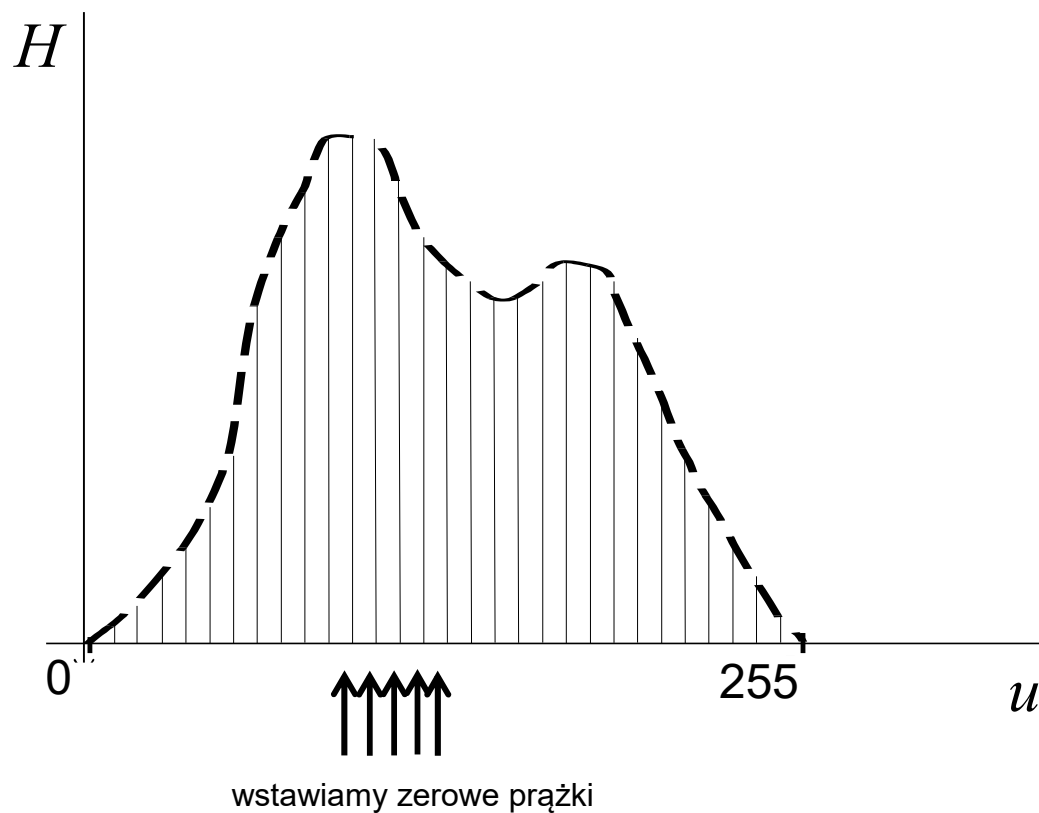
Poszerzanie zakresu dynamicznego

Przekształcenie



Poszerzanie zakresu dynamicznego (c.d.)

Histogram wynikowy

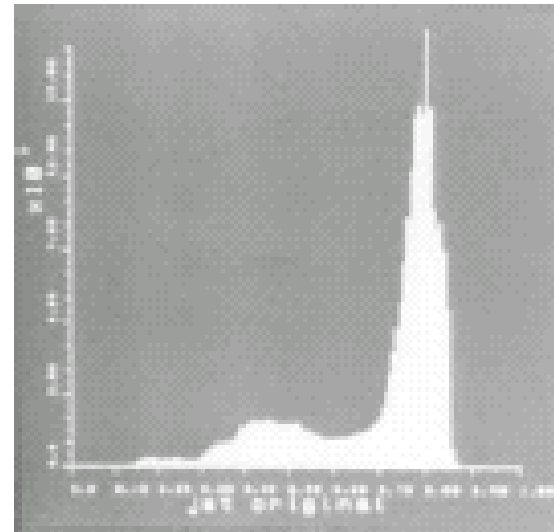


Za pomocą operacji punktowej poprawiliśmy **subiektywną** jakość obrazu.

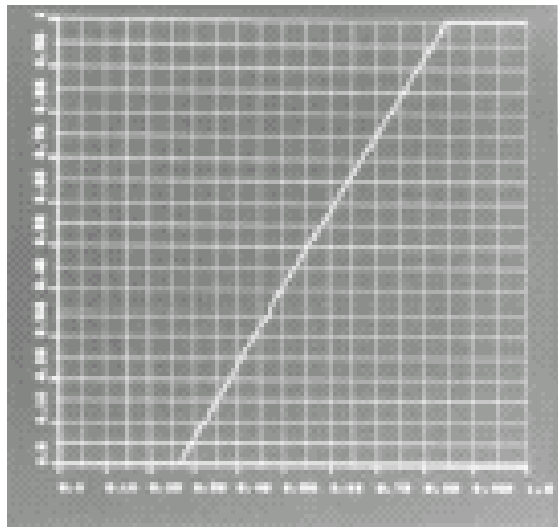
Poszerzanie zakresu dynamicznego



(a) Original



(b) Original histogram



Funkcja F opisująca przekształcenie



(d) Contrast stretched

Źródło: W.Pratt, Digital
Image Processing,
JWiley

Wyrównywanie histogramów

Histogram equalization

Cel : Uzyskanie obrazu, którego histogram $H(y)$ ma taką właściwość, że w poszczególnych przedziałach Δy suma wartości $H(y_i)$ jest stała

$$\begin{aligned} H(y_1) + H(y_1+1) + H(y_1+2) + \dots + H(y_1+\Delta y-1) &\cong \\ &\cong H(y_1+\Delta y) + H(y_1+\Delta y+1) + H(y_1+\Delta y+2) + \dots + H(y_1+2\Delta y-1) \cong \\ &\cong H(y_1+2\Delta y) + H(y_1+2\Delta y+1) + H(y_1+2\Delta y+2) + \dots + H(y_1+3\Delta y-1) \quad \text{itd.} \end{aligned}$$

Przekształcenie : $y = F(u) = D(u),$

gdzie $D(u)$ – dystrybuanta dla rozkładu

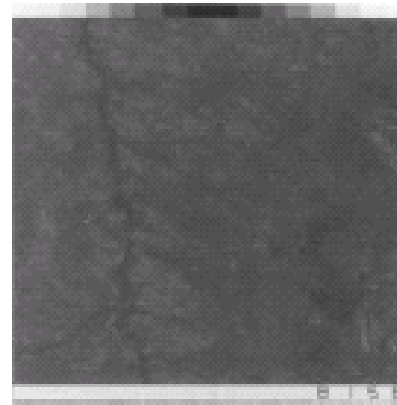
$p(u) = H(u)/N$, N - liczba punktów w obrazie

czyli $D(u) = \sum_{x \leq u} p(x)$.

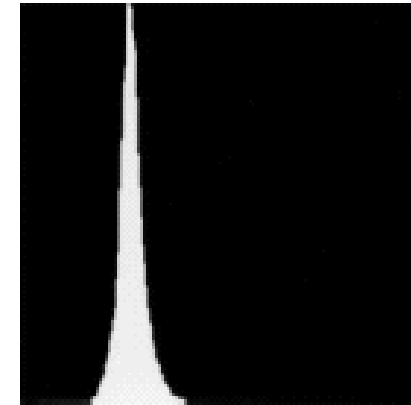
Porównanie wyrównywania histogramów z prostym poszerzaniem zakresu

Wynik poszerzania zakresu dynamicznego

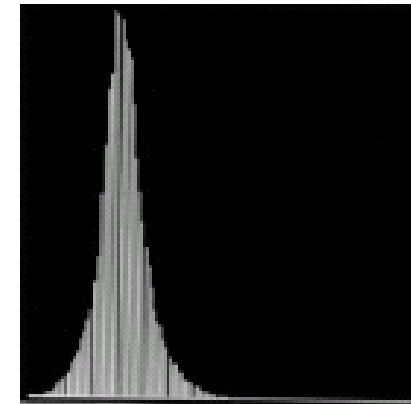
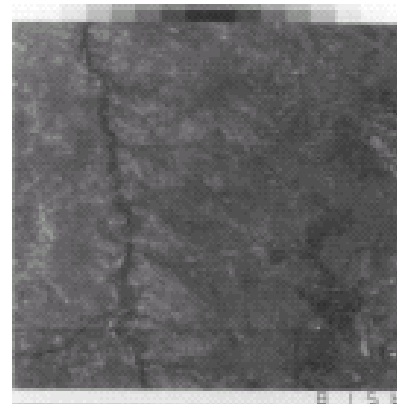
Wynik wyrównywania histogramu



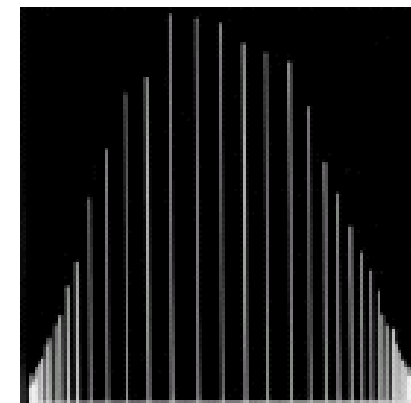
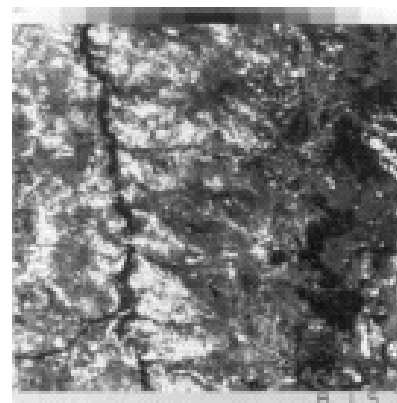
(a) Original



(b) Original histogram

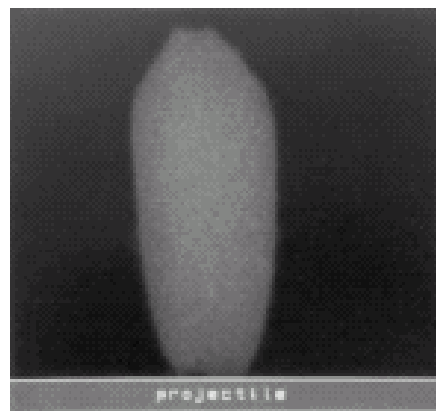


(d) Enhancement histogram

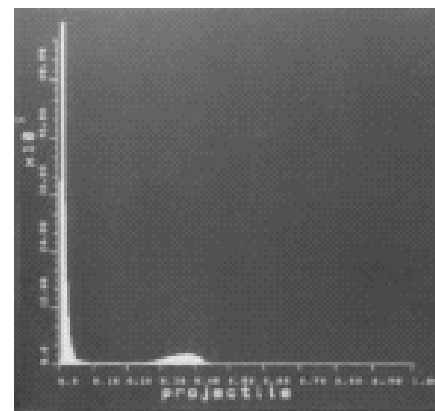


(f) Enhancement histogram

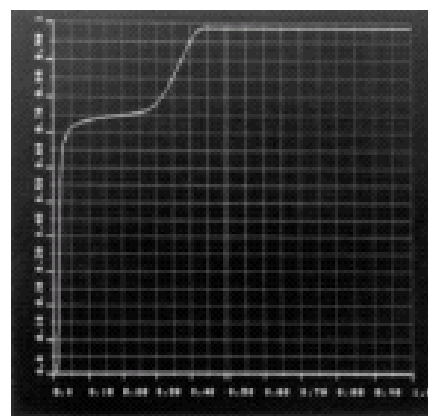
Wyrównywanie
histogramów
– jeszcze jeden
przykład



(a) Original



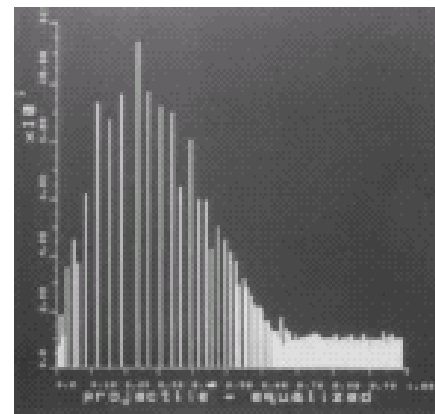
(b) Original histogram



Funkcja F opisująca przekształcenie



(d) Enhanced



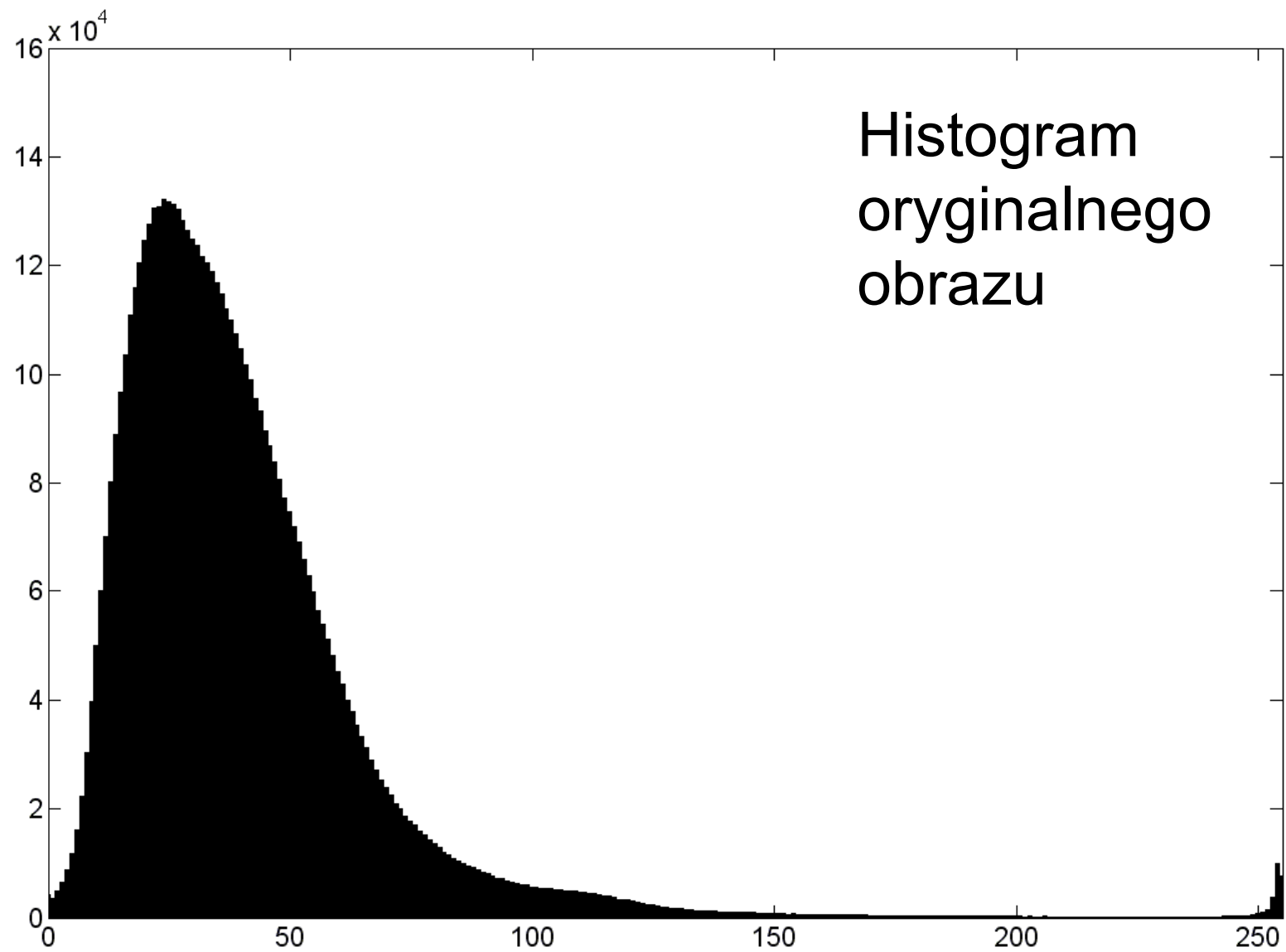
(e) Enhanced histogram

Oryginalny obraz

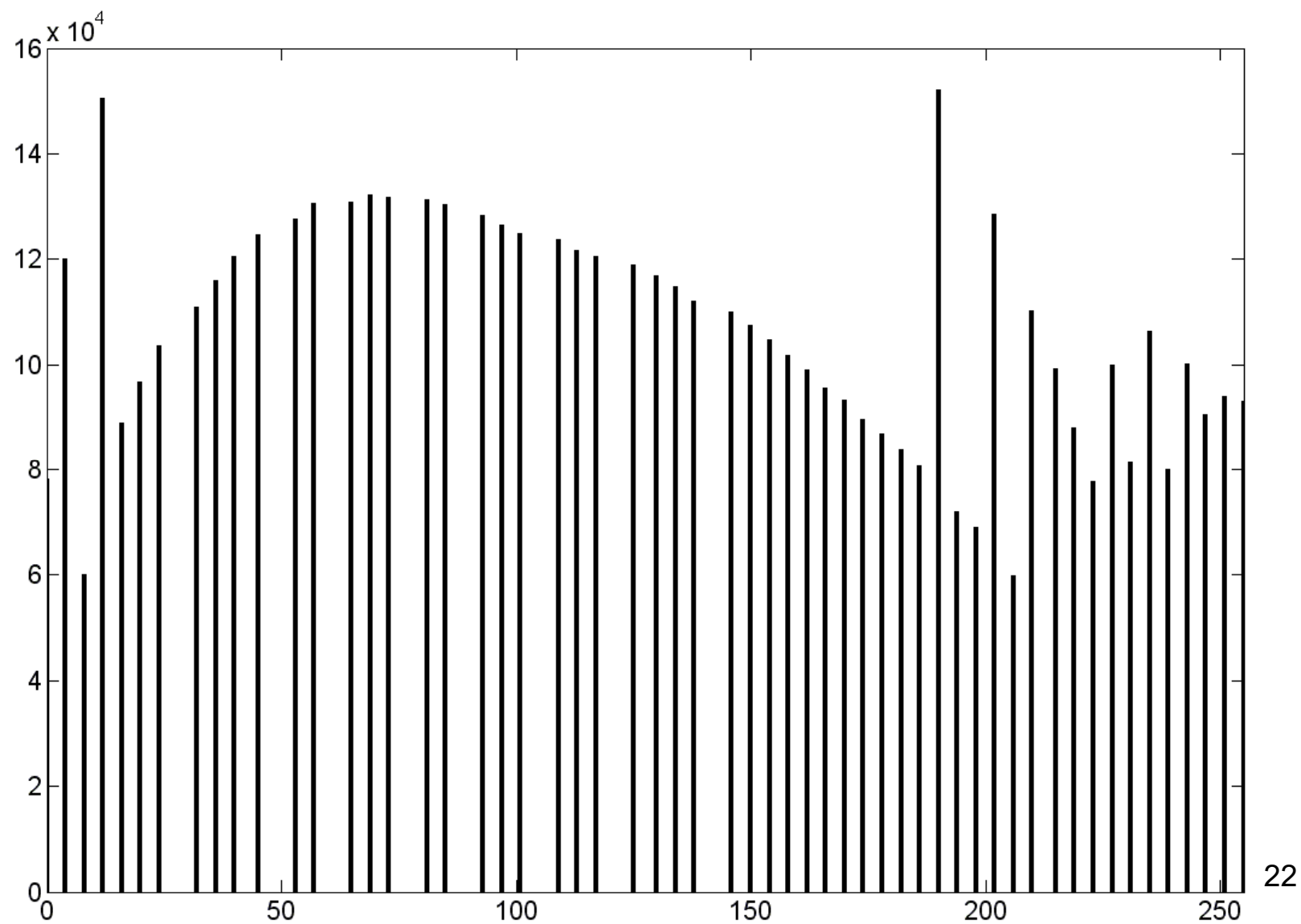
Słaby kontrast

Katedra w Sztokholmie





Histogram obrazu przetworzonego



Obraz przetworzony

Po wyrównaniu
histogamu



Przykład 3.4.1

- Dany jest obraz monochromatyczny o 8-bitowej reprezentacji próbek.
- Histogram tego obrazu jest następujący:
 $H(2) = 20\ 000$; $H(3) = 40\ 000$; $H(5) = 40\ 000$;
 $H(8) = 30\ 000$; $H(9) = 70\ 000$; $H(10) = 160\ 000$;
 $H(12) = 40\ 000$.
- 1) Wyznacz histogram obrazu o poszerzonym zakresie dynamicznym.
- 2) Wyznacz histogram obrazu po wyrównaniu histogramu.

3.4.2. Filtry liniowe FIR

Dwuwymiarowe filtry FIR - wprowadzenie

Filtry o skończonych odpowiedziach impulsowych

FIR – *f*inite *i*mpulse *r*esponse

$$y(n_1, n_2) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} h(k_1, k_2) \cdot u(n_1 - k_1, n_2 - k_2)$$

$u(n_1, n_2)$ – próbka sygnału (obrazu) wejściowego,

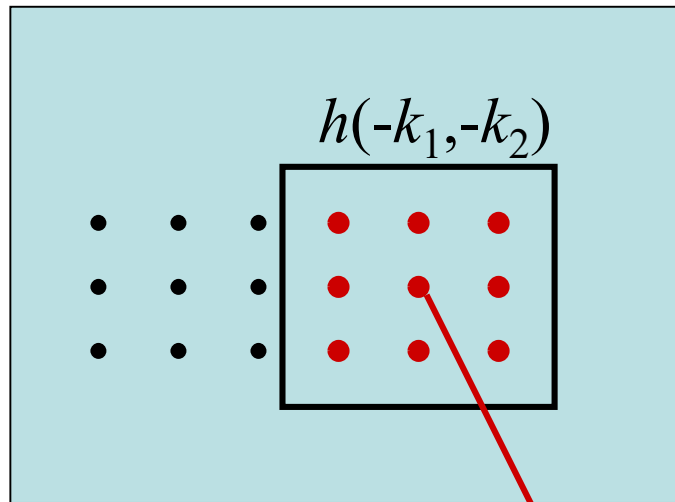
$y(n_1, n_2)$ – próbka sygnału (obrazu) wyjściowego,

$h(n_1, n_2)$ – próbka odpowiedzi impulsowej filtru.

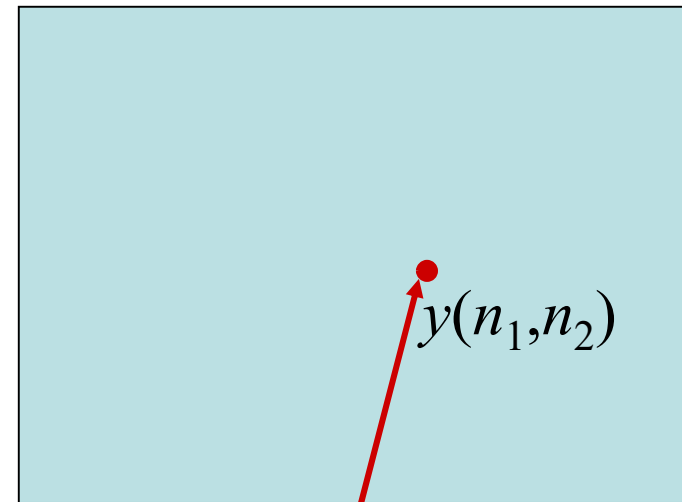
Próbki odpowiedzi impulsowej są różne od zera tylko w obszarze o rozmiarze $(2N_1+1) \times (2N_2+1)$.

Ilustracja działania filtru FIR

Obraz wejściowy



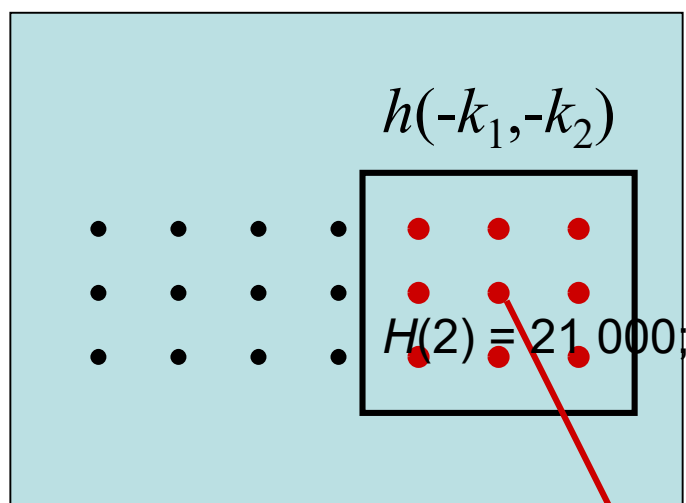
Obraz wyjściowy



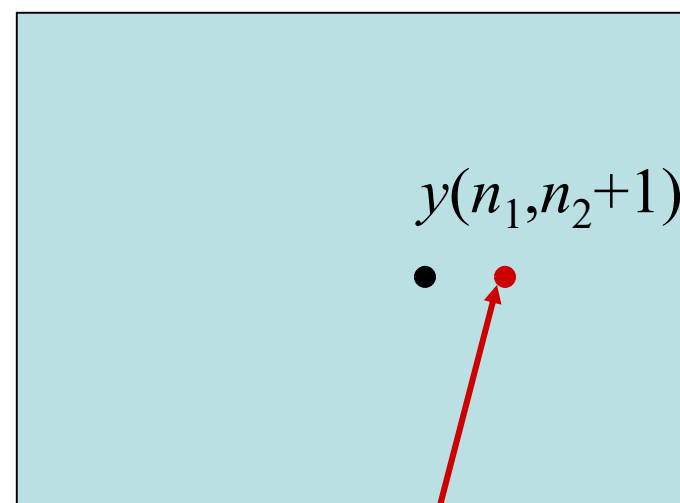
$$y(n_1, n_2) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} u(k_1, k_2) \cdot h(n_1 - k_1, n_2 - k_2) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} h(-k_1, -k_2) \cdot u(n_1 + k_1, n_2 + k_2)$$

Ilustracja działania filtru FIR

Obraz wejściowy



Obraz wyjściowy



$$y(n_1, n_2) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} u(k_1, k_2) \cdot h(n_1 - k_1, n_2 - k_2) = \sum_{k_1=-N_1}^{N_1} \sum_{k_2=-N_2}^{N_2} h(-k_1, -k_2) \cdot u(n_1 + k_1, n_2 + k_2)$$

Transmitancja dwuwymiarowego filtru FIR

$$H(j\omega_1, j\omega_2) = \mathcal{F} [h(n_1, n_2)]$$

$$H(j\omega_1, j\omega_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} h(n_1, n_2) \cdot \exp[-j\omega_1 n_1 - j\omega_2 n_2]$$

Ze względu dziedzinę niezerowej odpowiedzi impulsowej:

$$H(j\omega_1, j\omega_2) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) \cdot \exp[-j\omega_1 n_1 - j\omega_2 n_2]$$

Przykład

$$h = 0,2 \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Wyznaczyć transmitancję tego filtru.

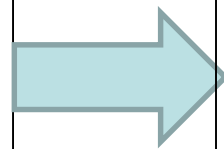
Symetrie filtrów FIR

$$h(n_1, n_2) = h(-n_1, n_2)$$

$$h(n_1, n_2) = h(n_1, -n_2)$$

$$H(j\omega_1, j\omega_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} h(n_1, n_2) \cdot \exp[-j\omega_1 n_1 - j\omega_2 n_2]$$

$$\begin{aligned} h(n_1, n_2) &= h(-n_1, n_2) \\ h(n_1, n_2) &= h(n_1, -n_2) \end{aligned}$$



Transmitancja ma
wartości rzeczywiste
- filtr „zerowej fazy”

Separowalność filtrów dwuwymiarowych

Filtr jest separowalny, jeżeli $h(n_1, n_2) = h_1(n_1) \cdot h_2(n_2)$

$$\text{Wtedy też } H(j\omega_1, j\omega_2) = H_1(j\omega_1) \cdot H_2(j\omega_2)$$

Przykład

$$h = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot [1 \quad 2 \quad 1] = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

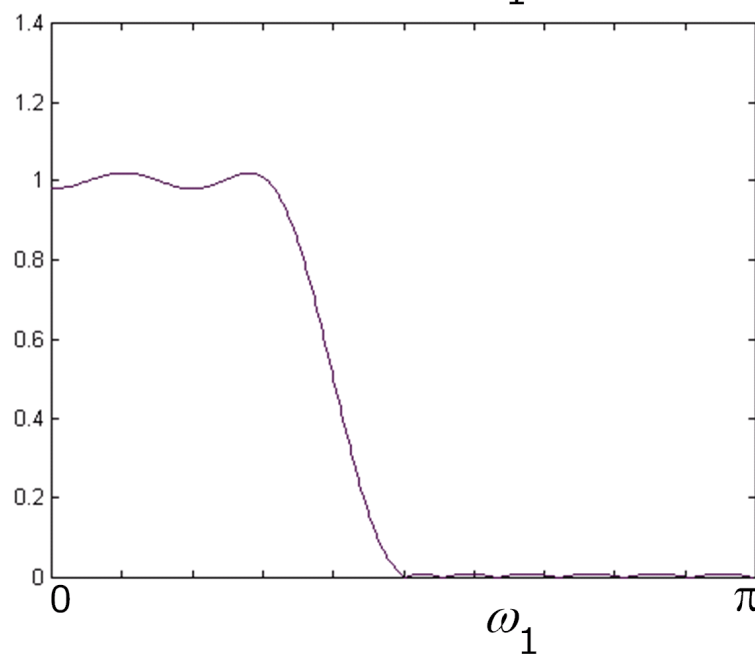
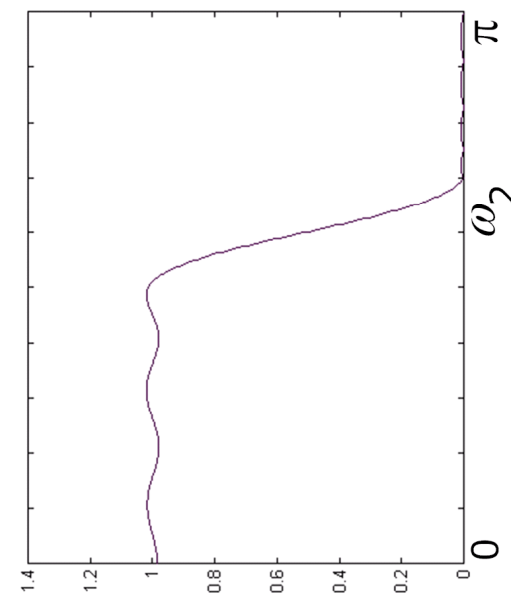
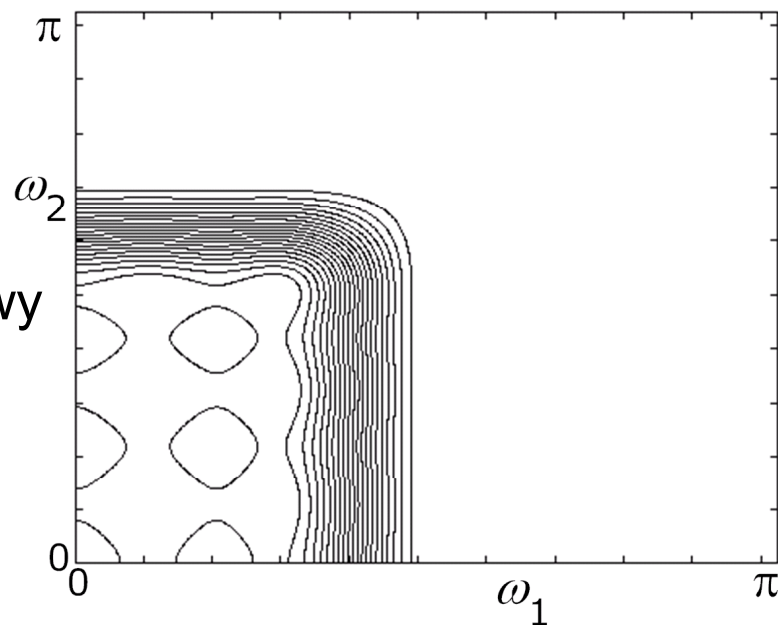
Filtr dwuwymiarowy można zrealizować przez złożenie dwóch jednowymiarowych filtrów działających w dwóch kierunkach.

Można w ten sposób zmniejszyć liczbę operacji

zamiast $(2N+1)^2$ mnożeń 2 $(2N+1)$ mnożeń.

Dwuwymiarowy filtr separowalny

Przykład :
Filtr
dolnoprzepustowy



Charakterystyka
amplitudowa filtru
o transmitancji $H_2(z_2)$

Charakterystyka amplitudowa
filtru o transmitancji $H_1(z_1)$

Proste dolnoprzepustowe filtry dwuwymiarowe

Postulaty: Zakładamy filtr zerowej fazy $h(n_1, n_2) = h(-n_1, n_2) = h(n_1, -n_2)$

Filtr nie zmienia składowej stałej

$$H(j\omega_1, j\omega_2) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) \cdot \exp[-j\omega_1 n_1 - j\omega_2 n_2]$$

$$H(j0, j0) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) \cdot \exp[-j0n_1 - j0n_2] = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) = 1$$

Filtry dolnoprzepustowe

$$H(j0, j0) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) = 1$$

Przykład prostego filtra

$$h = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

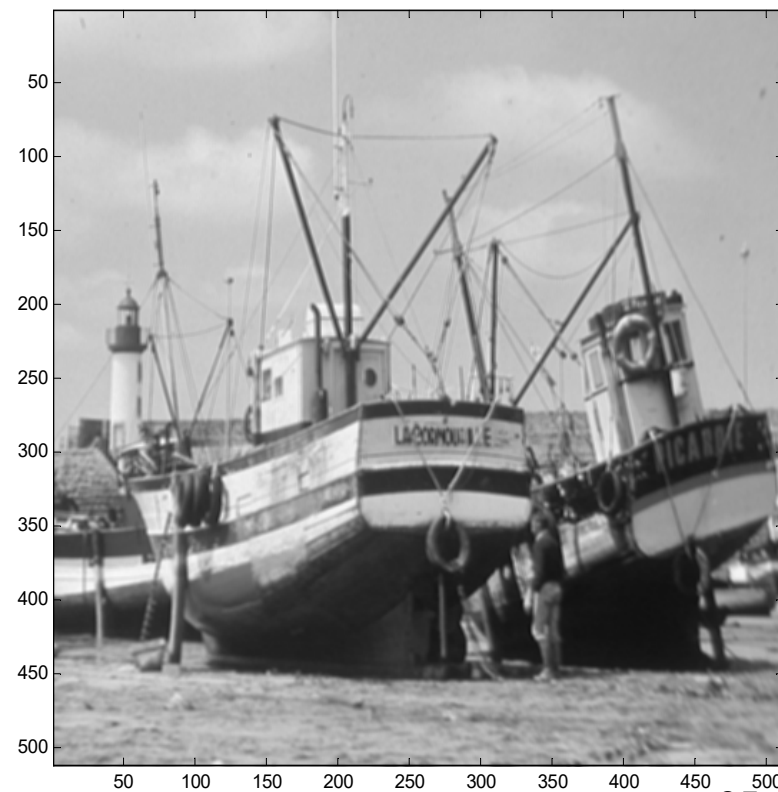
Przykład prostego filtra

$$h = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Obraz wejściowy



Obraz wyjściowy







Filtry dolnoprzepustowe

$$H(j0, j0) = \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} h(n_1, n_2) = 1$$

Filtr **decymacyjny** separowalny

$$h = (1/128) [-29 \quad 0 \quad 88 \quad 138 \quad 88 \quad 0 \quad -29]$$

Proste filtry górnoprzepustowe

Filtr **Sobela**

$h =$ -1 0 1

-2 0 2

-1 0 1

-1 -2 -1

0 0 0

1 2 1

Proste filtry górnoprzepustowe

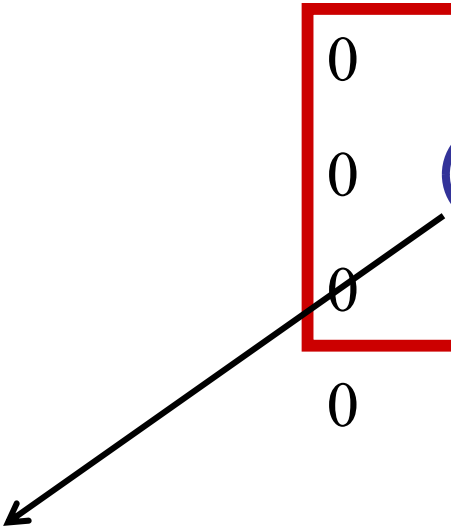
Filtr **Sobela**

$h =$ -1 0 1
-2 0 2
-1 0 1

obraz

0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100

0



Proste filtry górnoprzepustowe

Filtr **Sobela**

$$h = \begin{matrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{matrix}$$

obraz

0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100

0 **-400**

Proste filtry górnoprzepustowe

Filtr **Sobela**

$h =$ -1 0 1
-2 0 2
-1 0 1

obraz

0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	0	100	100

0 - 400 - 400

Proste filtry górnoprzepustowe

Filtr **Sobela**

$h =$ -1 0 1

-2 0 2

-1 0 1

obraz

0 0 0 100 100 100

0 0 0 100 100 100

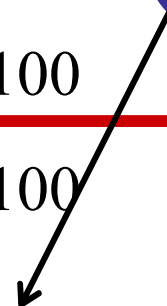
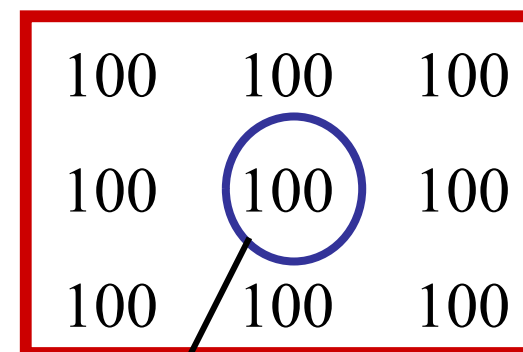
0 0 0 100 100 100

0 0 0 100 100 100

0 0 0 100 100 100

0 0 0 100 100 100

0 - 400 - 400 0



Filtry górnoprzepustowe – filtry Sobela



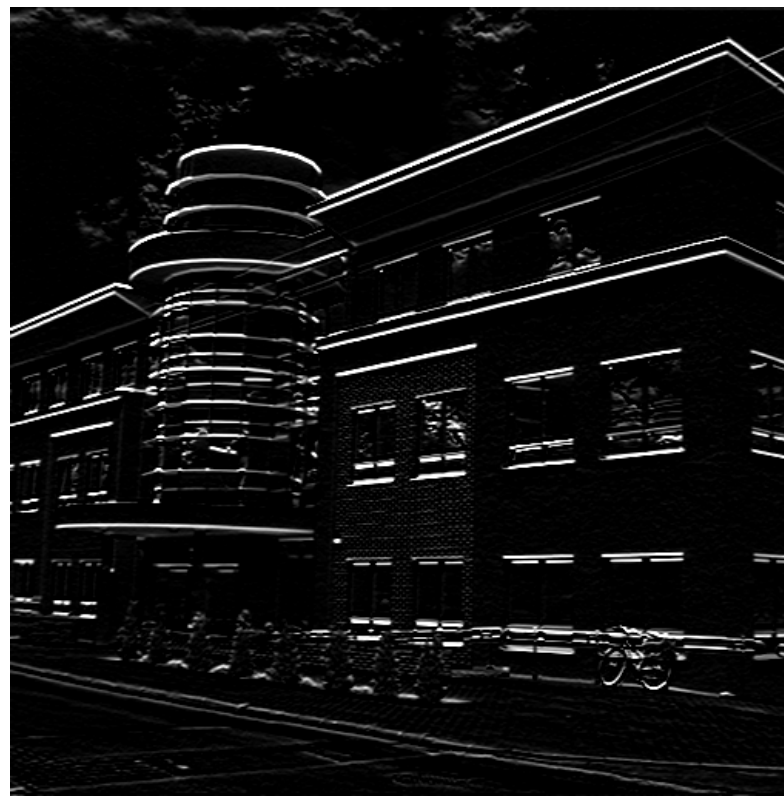
Filtry górnoprzepustowe – filtry Sobela



Filtry górnoprzepustowe – filtry Sobela



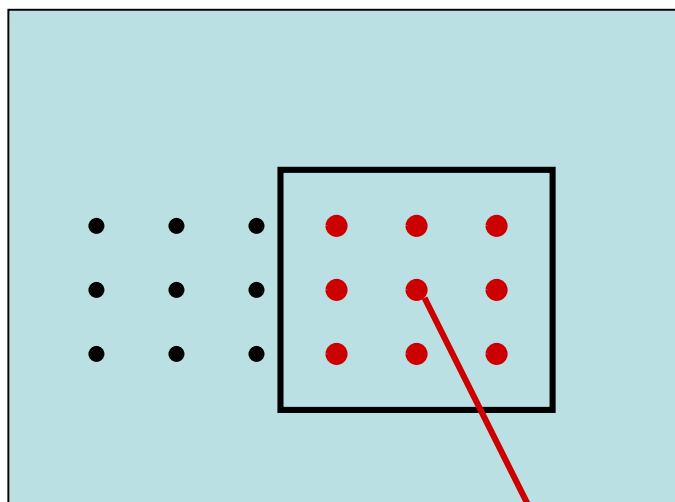
Filtry górnoprzepustowe – filtry Sobela



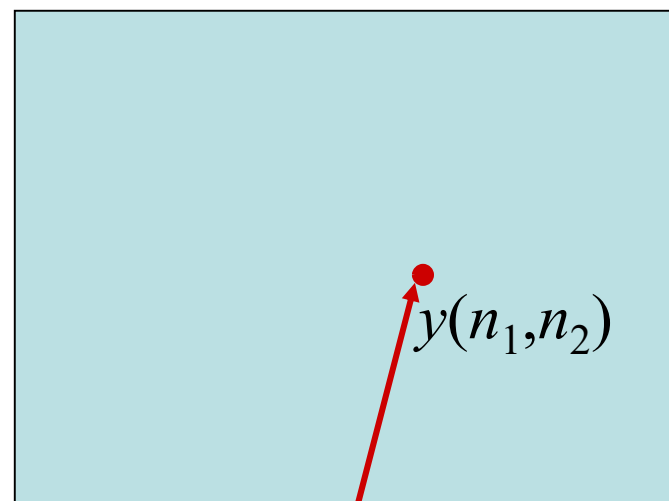
3.4.3. Filtry medianowe

Ilustracja działania filtru medianowego

Obraz wejściowy



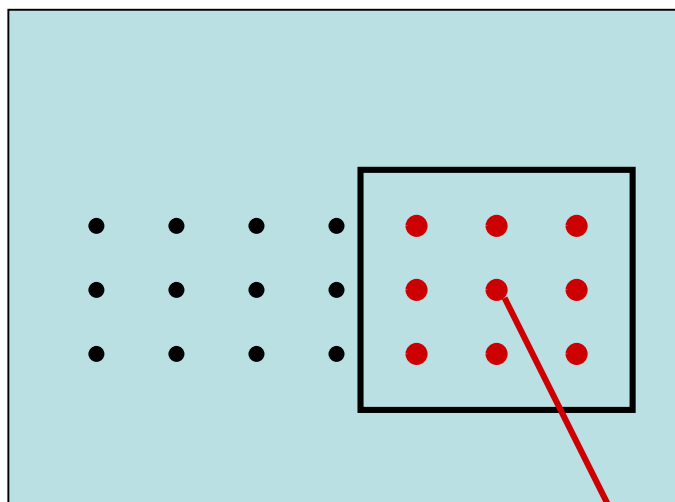
Obraz wyjściowy



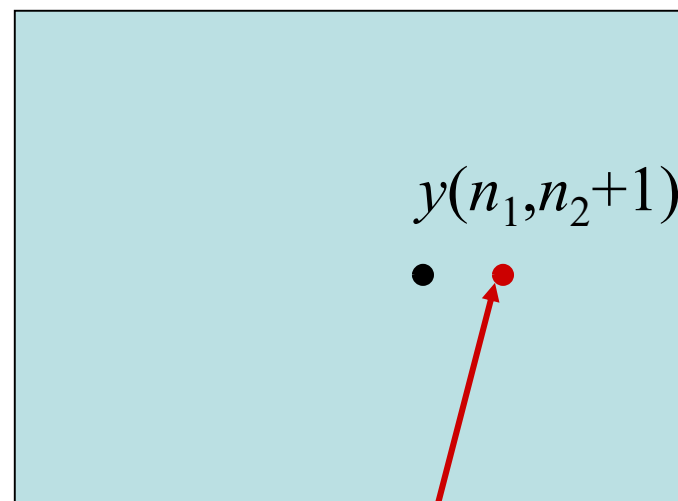
$$y(n_1, n_2) = \text{med} \{u(n_1 - k_1, n_2 - k_2); -N_1 \leq k_1 \leq N_1, -N_2 \leq k_2 \leq N_2\}$$

Ilustracja działania filtru medianowego

Obraz wejściowy



Obraz wyjściowy



$$y(n_1, n_2) = \text{med} \{u(n_1 - k_1, n_2 - k_2); -N_1 \leq k_1 \leq N_1, -N_2 \leq k_2 \leq N_2\}$$

Usuwanie szumów impulsowych za pomocą filtru medianowego

Oryginał



Obraz zaszumiony ($p=10\%$)



Usuwanie szumów impulsowych za pomocą filtru medianowego

Obraz przetworzony
filtrem o rozmiarze 3×3

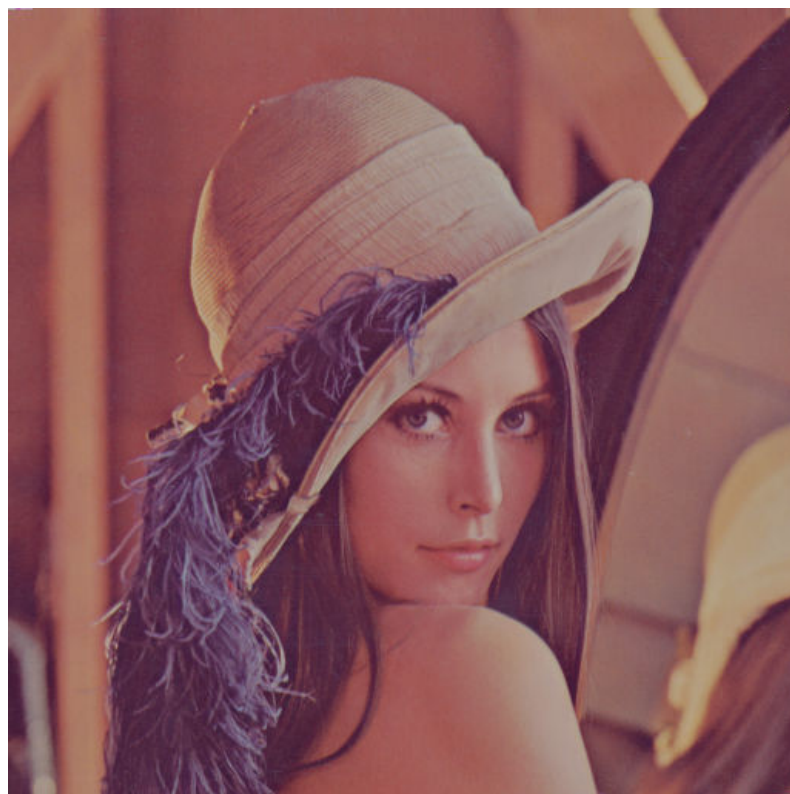


Obraz przetworzony
filtrem o rozmiarze 5×5

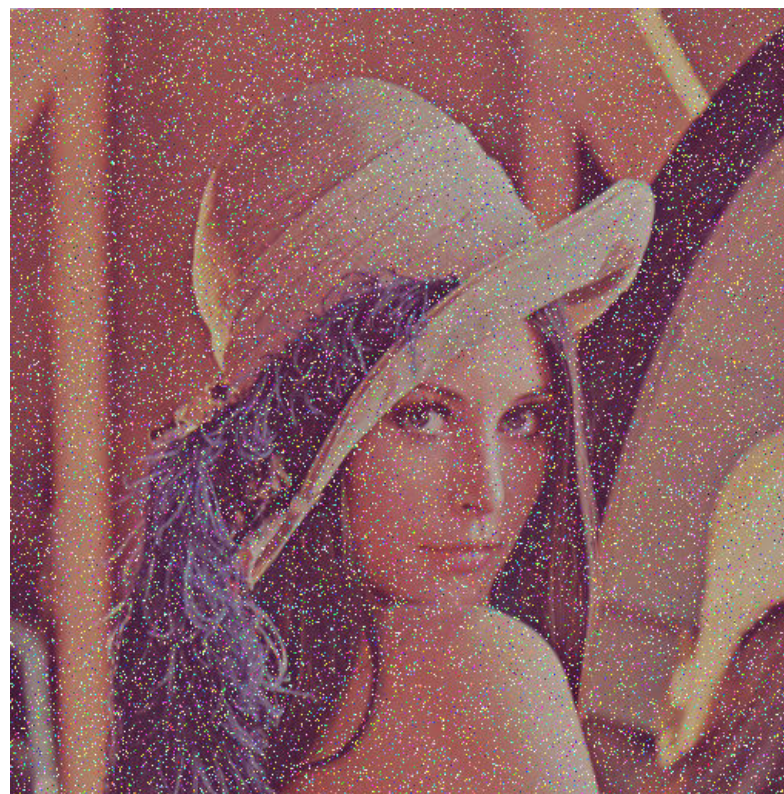


Usuwanie szumów impulsowych za pomocą filtru medianowego

Oryginał

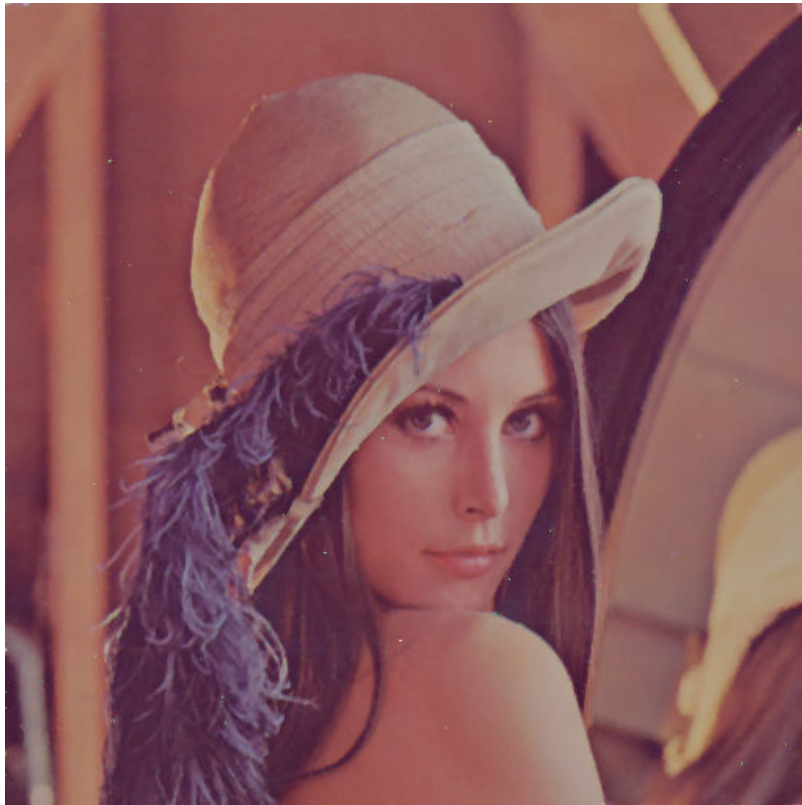


Obraz zaszumiony ($p=10\%$)

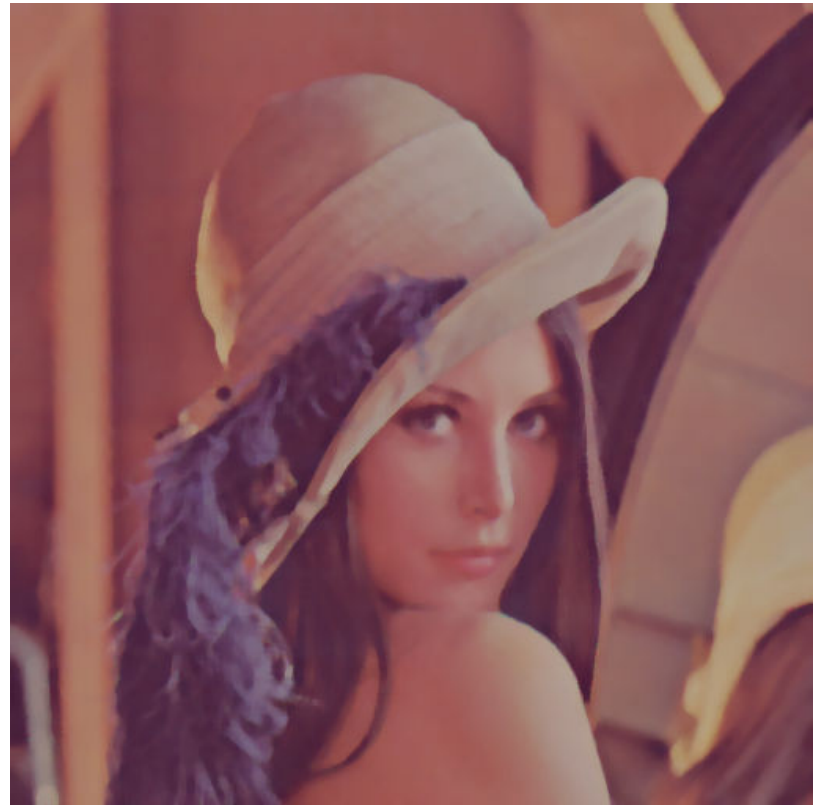


Usuwanie szumów impulsowych za pomocą filtru medianowego

Obraz przetworzony
filtrem o rozmiarze 3×3



Obraz przetworzony
filtrem o rozmiarze 5×5



3.4.4. Operacje morfologiczne

B – maska operacji (filtru)

Rozszerzanie (*dilation*)

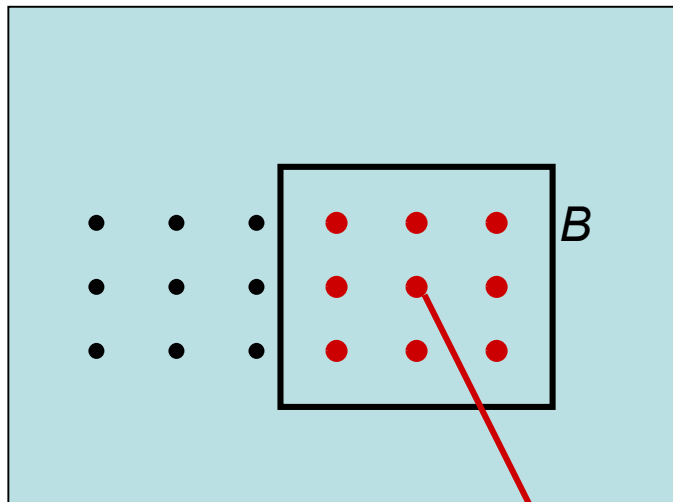
$$(f \oplus B)(x)$$

Ograniczanie (*erosion*)

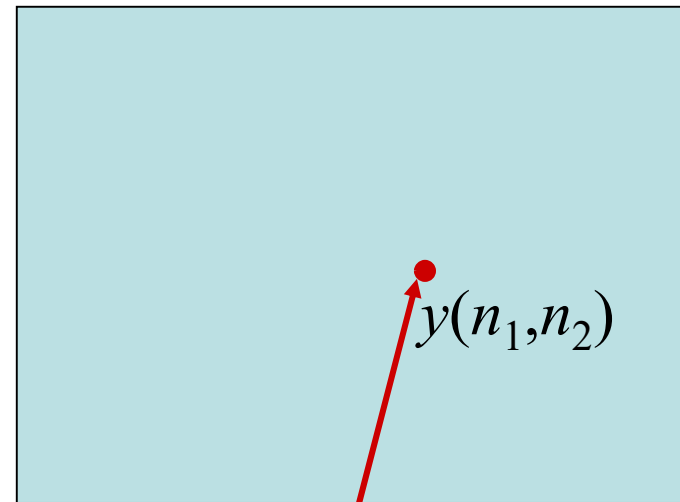
$$(f \ominus B)(x)$$

Ilustracja działania rozszerzania

Obraz wejściowy



Obraz wyjściowy



$$y = u \oplus B$$

$$y(n_1, n_2) = \max \{u(n_1 - k_1, n_2 - k_2); -N_1 \leq k_1 \leq N_1, -N_2 \leq k_2 \leq N_2\}$$

Oryginalny

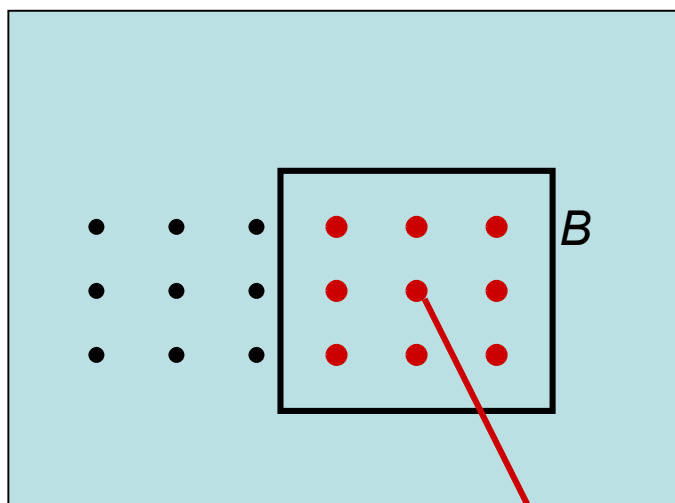


Filtr rozszerzający
5 x 5

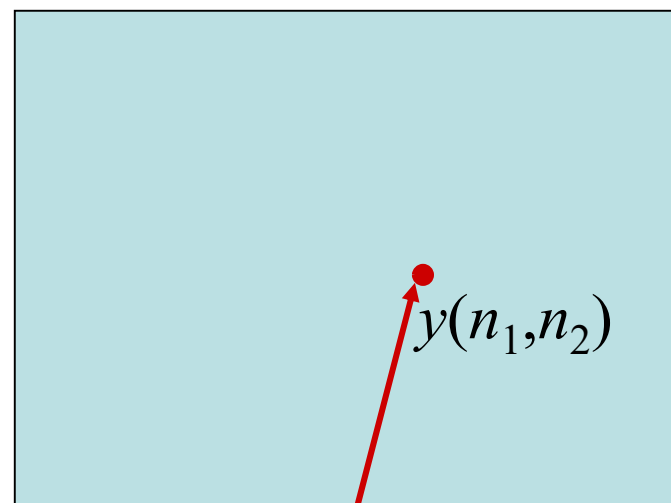


Ilustracja działania ograniczania

Obraz wejściowy



Obraz wyjściowy



$$y = u \ominus B$$

$$y(n_1, n_2) = \min \{u(n_1 - k_1, n_2 - k_2); -N_1 \leq k_1 \leq N_1, -N_2 \leq k_2 \leq N_2\}$$

Oryginalny



Filtr ograniczający
3 x 3



Filtr ograniczający
5 x 5



Przykład

obraz

0	0	0	100	100
0	0	0	100	100
0	0	0	100	100
0	0	50	100	100
0	50	50	100	100
50	50	50	100	100

1) Wyznaczyć odpowiedź
filtru medianowego
Maska: 3 X 3

2) Wyznaczyć
rozszerzenie i ograniczenie
Maska: 3 X 3

Gradient morfologiczny

$$G = (f \oplus B) - (f \ominus B)$$



Zadanie 1

- Częstotliwość próbkowania wynosi 100
Próbkowanie ma być jednowymiarowe

$$h = (1/128) [-29 \ 0 \ 88 \ 138 \ 88 \ 0 \ -29] .$$

- Wyznaczyć charakterystykę amplitudową
filtru dla częstotliwości przestrzennych 10 i
40 próbek na cal.

Zadanie 2

- Częstotliwość próbkowania wynosi 100 próbek na cal w obu kierunkach .

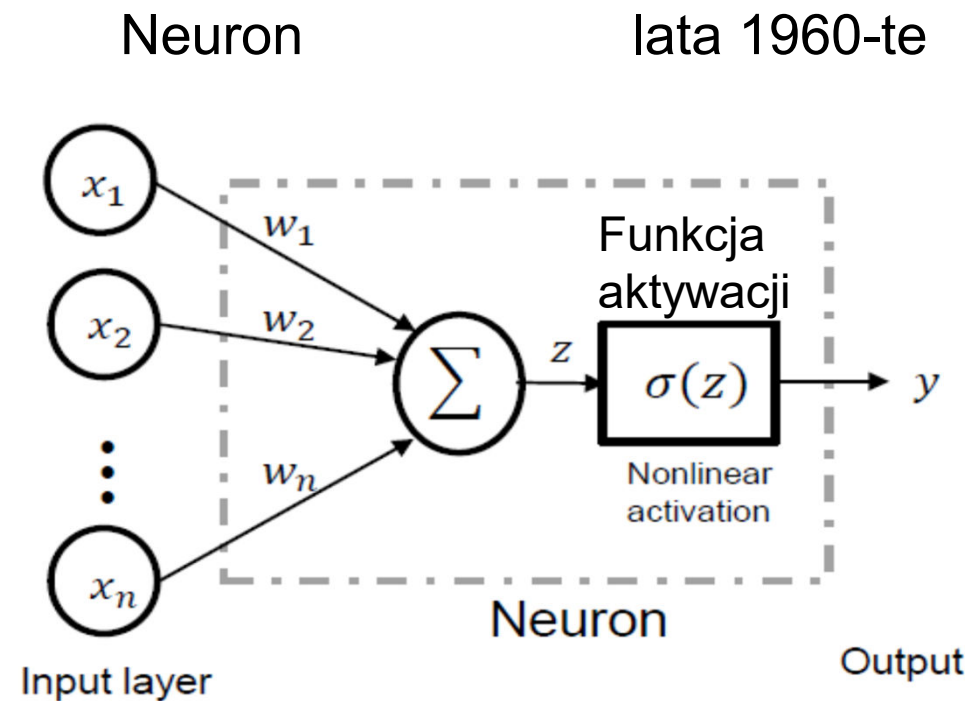
Filtr decymacyjny separowalny złożony z dwóch filtrów jednowymiarowych

$$h = (1/128) [-29 \quad 0 \quad 88 \quad 138 \quad 88 \quad 0 \quad -29] .$$

- Dla filtru dwuwymiarowego złożonego z dwóch filtrów jednowymiarowych zdefiniowanych jak wyżej wyznaczyć charakterystykę amplitudową dla częstotliwości przestrzennych (10, 10) i (40,40) próbek na cal.

3.5. Wykorzystanie sztucznych sieci neuronowych w analizie obrazów

Pierwsze sztuczne sieci neuronowe



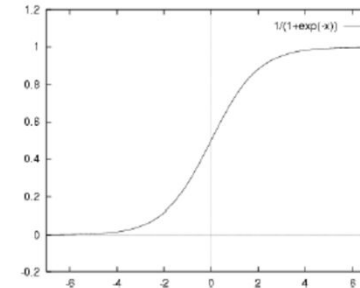
Dawniejsze funkcje aktywacji

Activation function should be differentiable so as to perform to compute gradient of output error (loss function) with respect to unknown weights.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

- **Logistic function (sigmoid)**

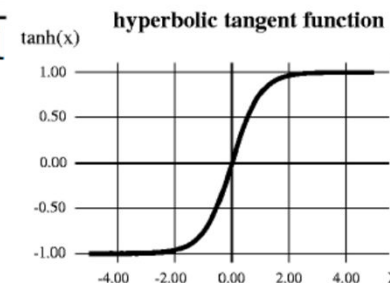
- Vanishing gradients
- Sigmoids saturate and kill gradients.
- Output is between 0 and 1, not 0-centered



$$\sigma(z) = \frac{e^{2z} - 1}{e^{2z} + 1}$$

- **Hyperbolic tangent**

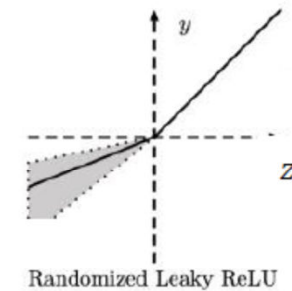
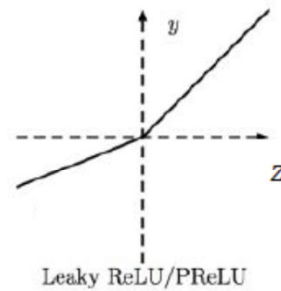
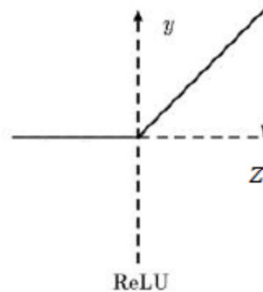
- Output is 0-centered in between -1 to 1
- Vanishing gradients



Source: M. Tekalp, ICIP 2019

Obeenie często stosowane funkcje aktywacji

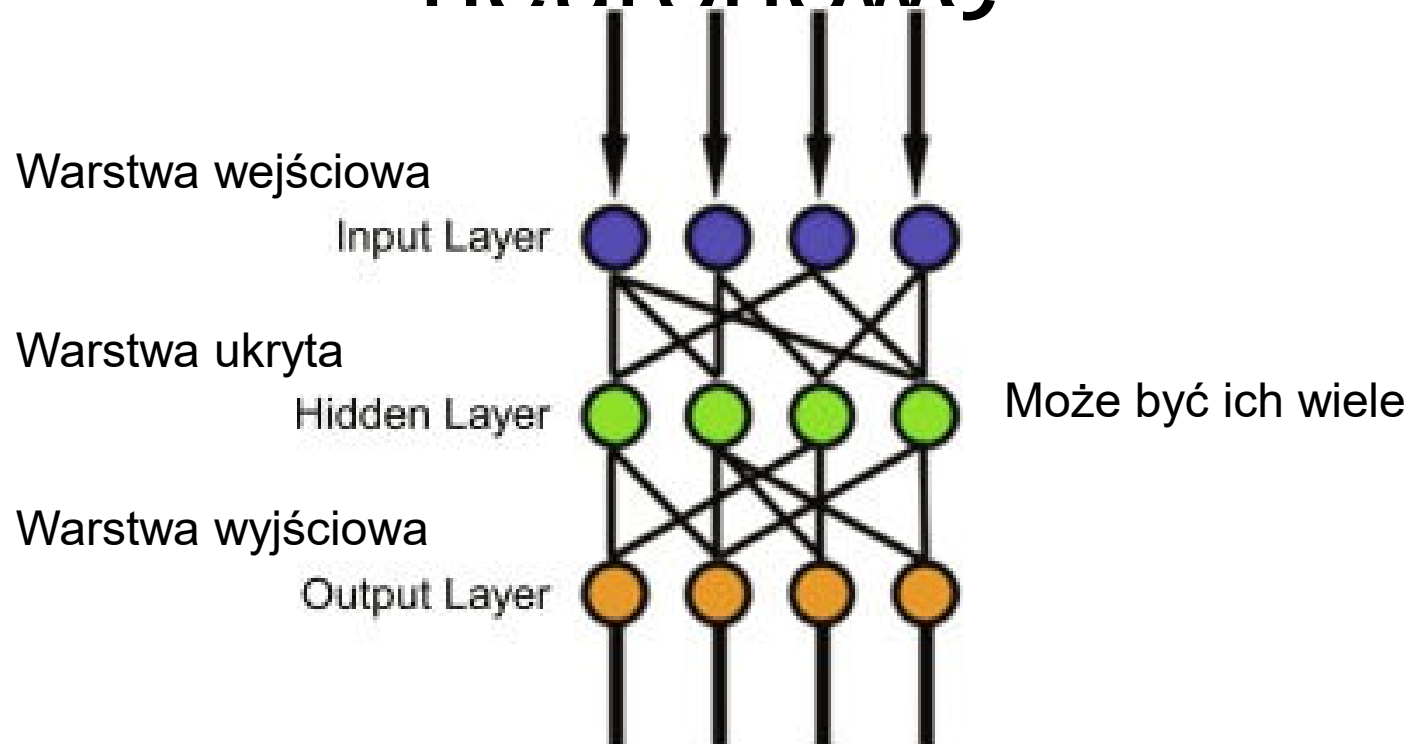
- Avoid **vanishing gradients problem**
- Rectified linear unit (ReLU) $\sigma(z) = \max\{0, z\}$



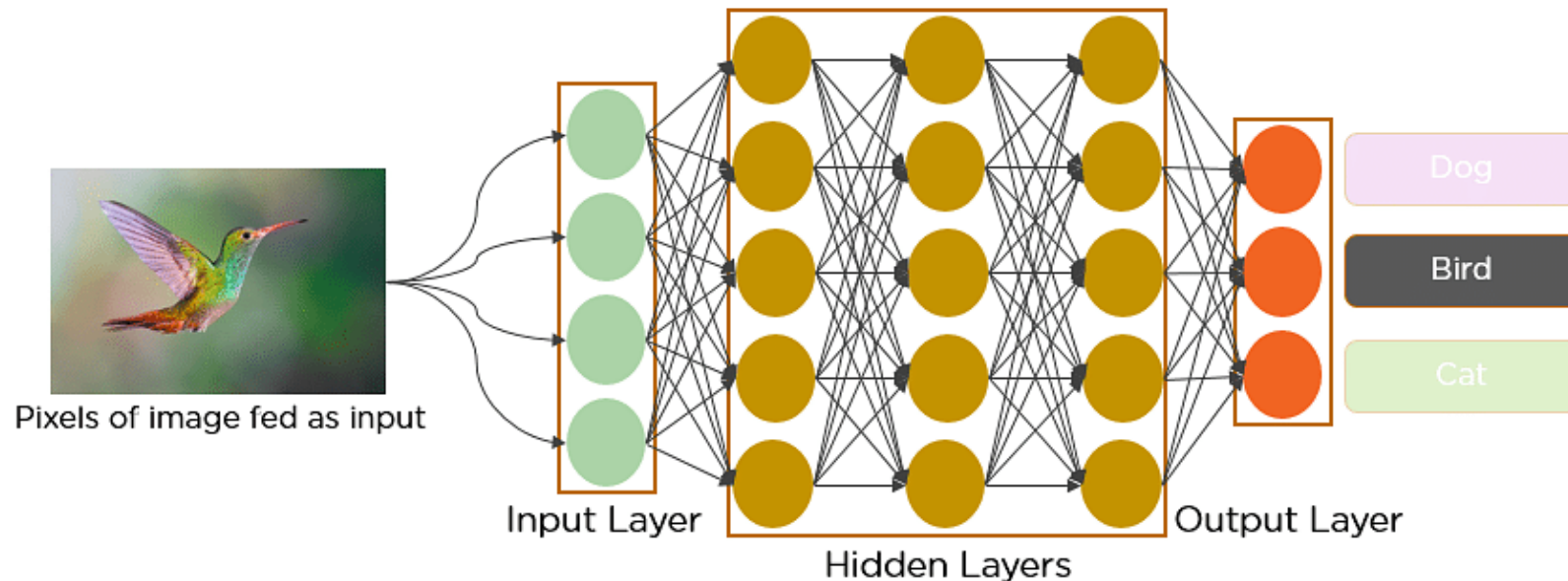
- Only used in hidden layers – in the output layer use softmax for classification problems and linear layer for regression problems
- **Dead neuron problem** – use Leaky ReLU

Source: M. Tekalp, ICIP 2019

Wielowarstwowe sieci neuronowe



Przykład zastosowania sieci wielowarstwowej do klasyfikacji obrazów



Ogólne twierdzenie o aproksymacji

Universal Approximation Theorem

- Dla dowolnej ciągłej funkcji $f(x)$ można znaleźć wielowarstwową sieć neuronową o wyjściu określonym funkcją $g(x)$, dla której

$$|f(x) - g(x)| < \varepsilon$$

dla wszystkich wejść oraz określonego maksymalnego błędu aproksymacji ε .

Głęboka sieć neuronowa (*deep neural network*)

- Modelowana przez wielokrotne złożenie prostych sieć $f(\mathbf{x}) = f_1 \left(f_2 \left(\dots \left(f_n(\mathbf{x}) \right) \right) \right)$

Uczenie nadzorowane (*supervised training*)

Dane pary wejście-wyjście: $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}), i = 1, \dots, N$.

Problem: Znaleźć wektor wag $\mathbf{w}$.

Minimalizacja kosztu, np.: $\sum_{i=1}^N (\mathbf{y}^{(i)} - \hat{\mathbf{y}}(\mathbf{w}, \mathbf{x}^{(i)}))^2$.

Rozwiązanie: metoda gradientu prostego, propagacja wsteczna.

Back-propagation – Propagacja wsteczna

- Network is initiated with random weights.
- Forward pass: Given input, the output error is computed.
- Backward pass: The gradient of the output error function with respect to weights is computed to update previous weights
 - Chain rule of differentiation
 - Shared computations among layers -
A computational graph to reuse partials

Source: M. Tekalp, ICIP 2019

Przykład propagacji wstecznej – krok wprzód

- Forward pass:

$$net_{h1} = w_1 i_1 + w_2 i_2 + b_1$$

$$out_{h1} = \frac{1}{1 + e^{-net_{h1}}}$$

$$net_{h2} = w_3 i_1 + w_4 i_2 + b_1$$

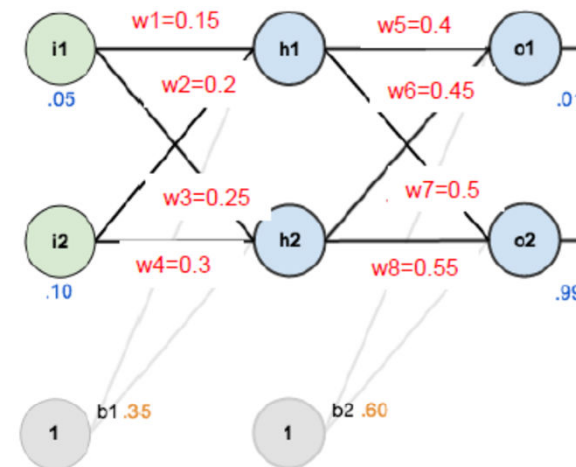
$$out_{h2} = \frac{1}{1 + e^{-net_{h2}}}$$

$$net_{o1} = w_5 out_{h1} + w_6 out_{h2} + b_2$$

$$out_{o1} = \frac{1}{1 + e^{-net_{o1}}}$$

$$net_{o2} = w_7 out_{h1} + w_8 out_{h2} + b_2$$

$$out_{o2} = \frac{1}{1 + e^{-net_{o2}}}$$



$$\begin{aligned} net_{h1} &= 0.3775 \\ out_{h1} &= 0.593269992 \\ net_{h2} &= 0.3925 \\ out_{h2} &= 0.596884378 \end{aligned}$$

$$\begin{aligned} net_{o1} &= 1.105905967 \\ out_{o1} &= 0.75136507 \\ net_{o2} &= 1.2249207 \\ out_{o2} &= 0.772928465 \end{aligned}$$

Przykład propagacji wstecznej – krok wstecz

- Backward pass:

- Output Layer

$$E_{total} = \frac{1}{2}(0.01 - out_{o1})^2 + \frac{1}{2}(0.99 - out_{o2})^2$$

$$\frac{\delta E_{total}}{\delta w_5} = \frac{\delta E_{total}}{\delta out_{o1}} \frac{\delta out_{o1}}{\delta net_{o1}} \frac{\delta net_{o1}}{\delta w_5}$$

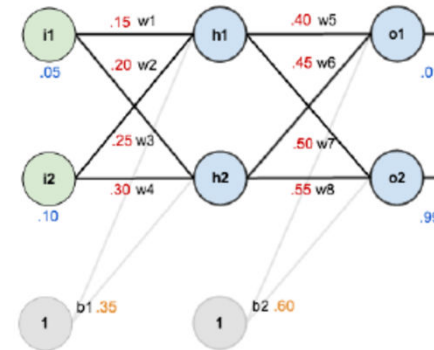
$$\frac{\delta E_{total}}{\delta out_{o1}} = -(0.01 - out_{o1})$$

$$\frac{\delta out_{o1}}{\delta net_{o1}} = \frac{e^{-net_{o1}}}{(1+e^{-net_{o1}})^2} = \frac{1}{1+e^{-net_{o1}}} \frac{e^{-net_{o1}}}{1+e^{-net_{o1}}} = out_{o1}(1 - out_{o1})$$

$$\frac{\delta net_{o1}}{\delta w_5} = out_{h1}$$

- Gradient-descent update

$$w_5 = w_5 + \alpha \frac{\delta E_{total}}{\delta w_5} \quad w_6 = w_6 + \alpha \frac{\delta E_{total}}{\delta w_6} \quad w_7 = w_7 + \alpha \frac{\delta E_{total}}{\delta w_7} \quad w_8 = w_8 + \alpha \frac{\delta E_{total}}{\delta w_8}$$



$$\frac{\delta E_{total}}{\delta w_5} = (0.74136)(0.18681)(0.59326) = 0.08216$$

Compute

$$\frac{\delta E_{total}}{\delta w_6}, \frac{\delta E_{total}}{\delta w_7}, \frac{\delta E_{total}}{\delta w_8}$$

similarly

Przykład propagacji wstecznej – krok wstecz (c.d.)

- Backward pass:

- Hidden Layer

$$\frac{\delta E_{total}}{\delta w_1} = \frac{\delta E_{total}}{\delta out_{h1}} \frac{\delta out_{h1}}{\delta net_{h1}} \frac{\delta net_{h1}}{\delta w_1}$$

$$\frac{\delta net_{h1}}{\delta w_1} = i_1$$

$$\frac{\delta out_{h1}}{\delta net_{h1}} = \frac{1}{1+e^{-net_{h1}}} \frac{e^{-net_{h1}}}{1+e^{-net_{h1}}} = out_{h1}(1-out_{h1})$$

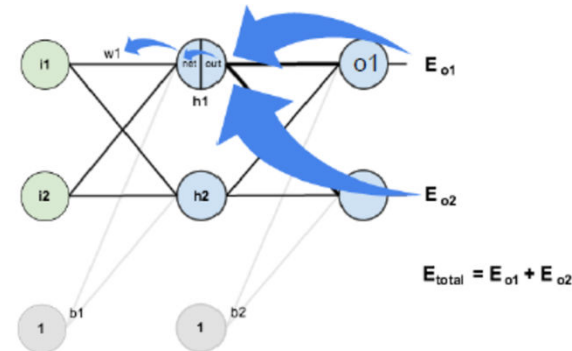
$$\frac{\delta E_{total}}{\delta out_{h1}} = \frac{\delta E_{o1}}{\delta out_{h1}} + \frac{\delta E_{o2}}{\delta out_{h1}}$$

$$\frac{\delta E_{o1}}{\delta out_{h1}} = \frac{\delta E_{o1}}{\delta net_{o1}} \frac{\delta net_{o1}}{\delta out_{h1}}$$

$$\frac{\delta E_{o1}}{\delta net_{o1}} = \frac{\delta E_{o1}}{\delta out_{o1}} \frac{\delta out_{o1}}{\delta net_{o1}}, \quad \frac{\delta net_{o1}}{\delta out_{h1}} = w_5$$

- Gradient-descent update

$$w_1 = w_1 + \alpha \frac{\delta E_{total}}{\delta w_1} \quad w_2 = w_2 + \alpha \frac{\delta E_{total}}{\delta w_2} \quad w_3 = w_3 + \alpha \frac{\delta E_{total}}{\delta w_3} \quad w_4 = w_4 + \alpha \frac{\delta E_{total}}{\delta w_4}$$



$$E_{o1} = \frac{1}{2} (0.01 - out_{o1})^2$$

$$E_{o2} = \frac{1}{2} (0.99 - out_{o2})^2$$

$$\frac{\delta E_{total}}{\delta w_1} = (0.036)(0.241)(0.05) = 0.000438$$

Compute

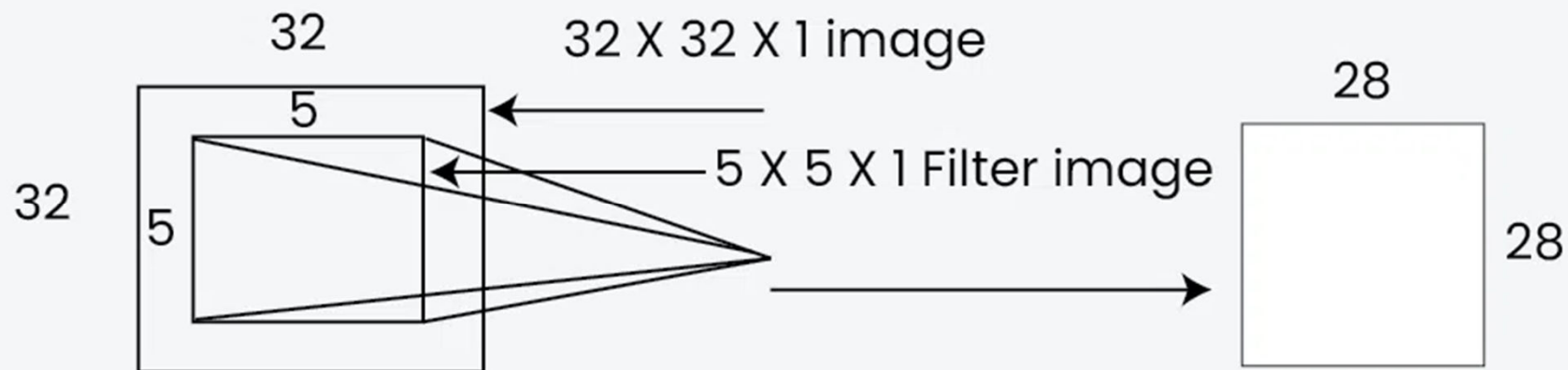
$$\frac{\delta E_{total}}{\delta w_2}, \frac{\delta E_{total}}{\delta w_3}, \frac{\delta E_{total}}{\delta w_4}$$

similarly

<https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/>

Sieci splotowe (*convolutional networks*)

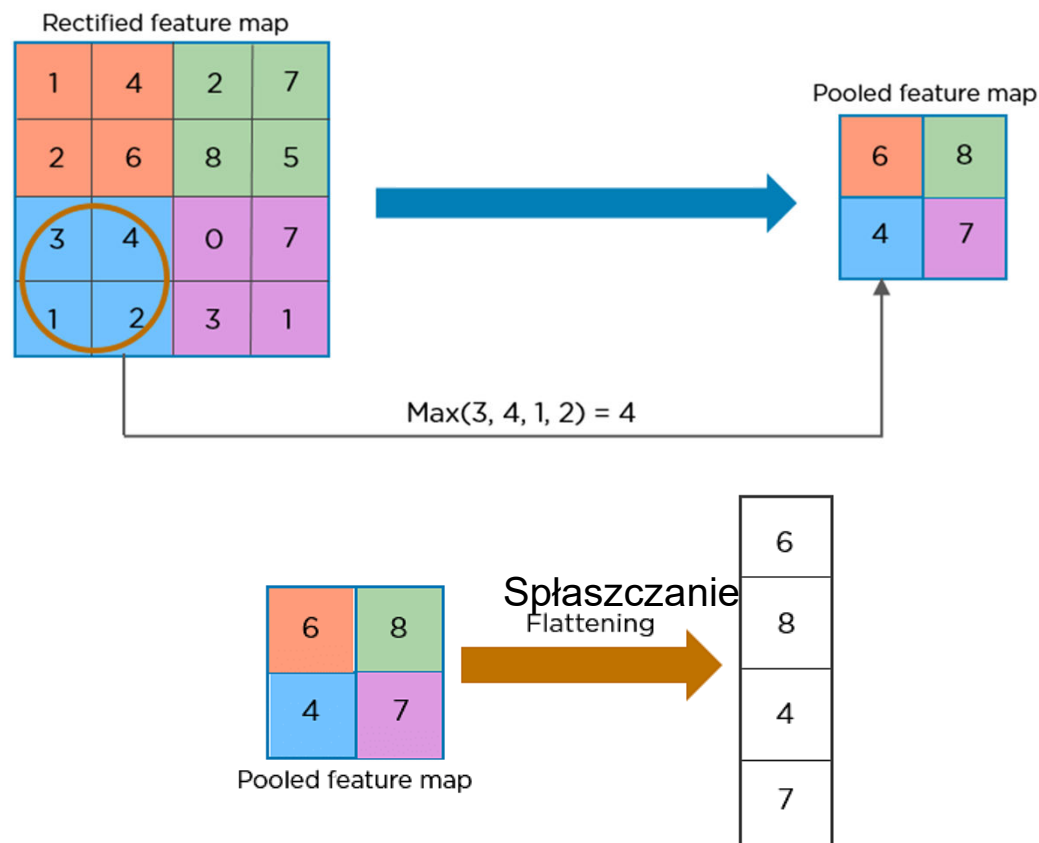
Warstwa splotowa Convolution layer



Source: <https://www.geeksforgeeks.org/>

Sieci splotowe (*convolutional networks*)

Feature map
= Mapa cech



<https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network>

Sieci splotowe (*convolutional networks*)

Warstwa splotowa

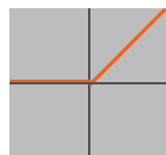
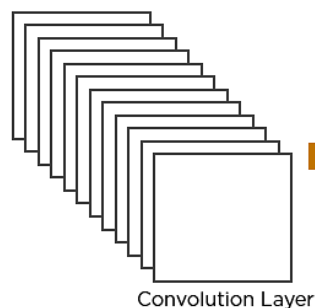
Próbki obrazu

Filtr

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

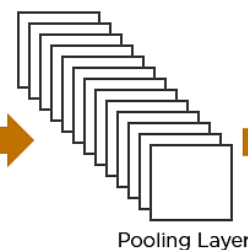
Input Image

Convolution



ReLU

Pooling



Flattening



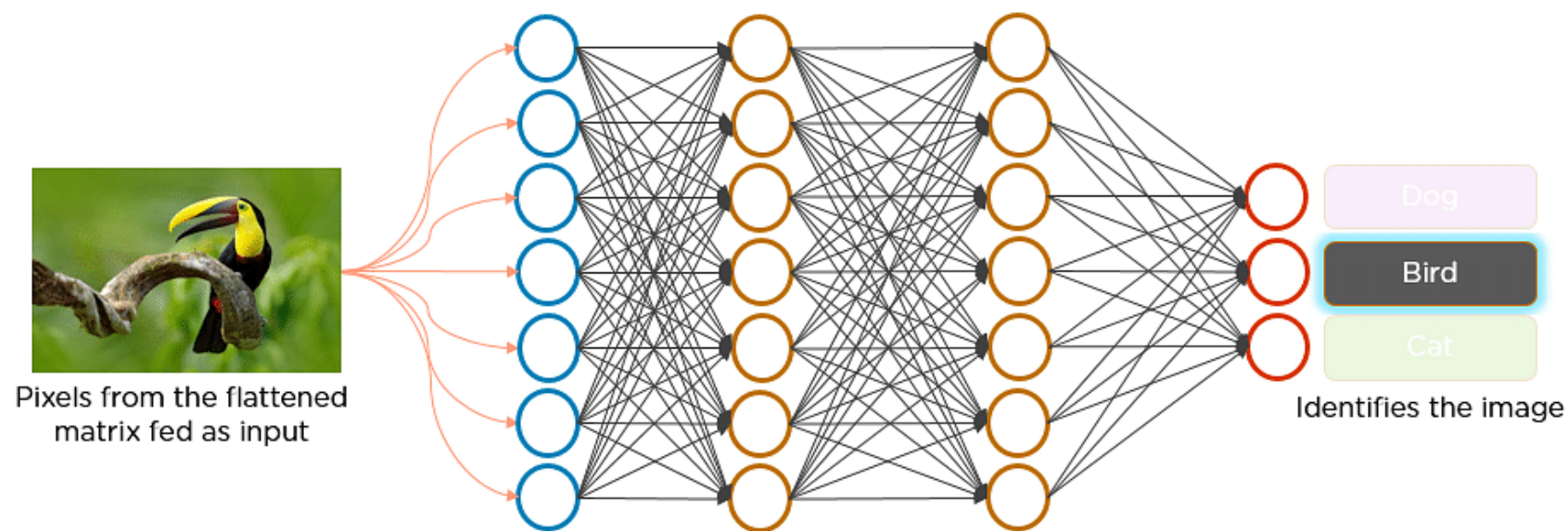
Input to the to final layer

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

1	0	1
0	1	0
1	0	1

Sieci splotowe (*convolutional networks*)

Klasyfikacja



<https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network>

Wykrywanie obiektu – *Object detection*

Wykrywanie obiektów (Object detection):

- Na wejściu: obraz.
- Na wyjściu:
 - ramki ograniczające dla każdego obiektu wraz z ich lokalizacją (np. współrzędne x, y);
 - etykieta klasy obiektu dla każdej ramki ograniczającej;
 - prawdopodobieństwo poprawności określenia klasy dla każdej ramki ograniczającej.

Wykrywanie obiektu – *Object detection*

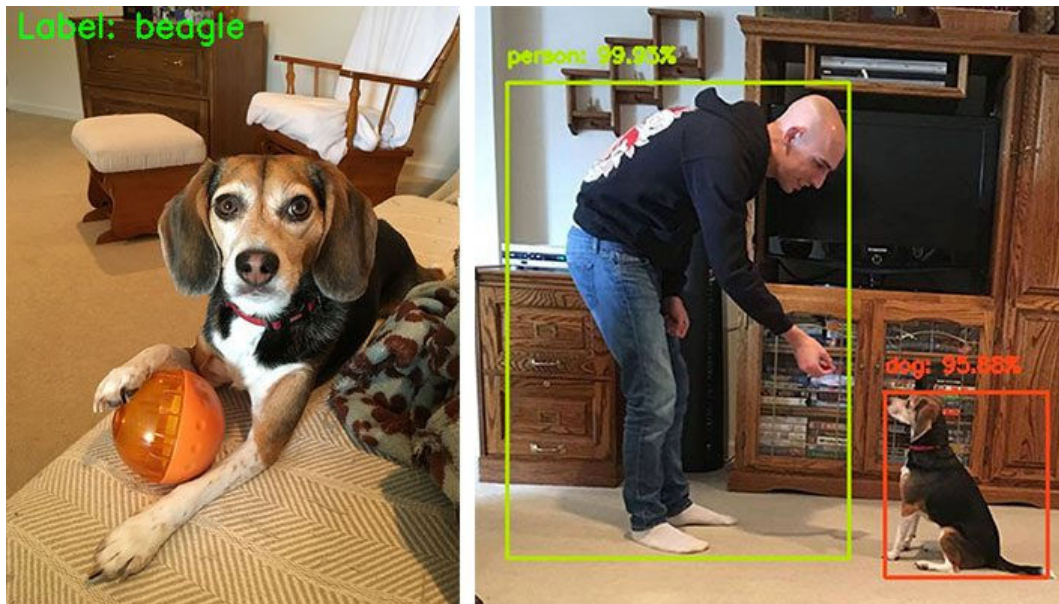
Należy odróżniać wykrywanie obiektów (*object detection*) od klasyfikacji obrazu (*image classification*).

Klasyfikacja obrazu (image classification):

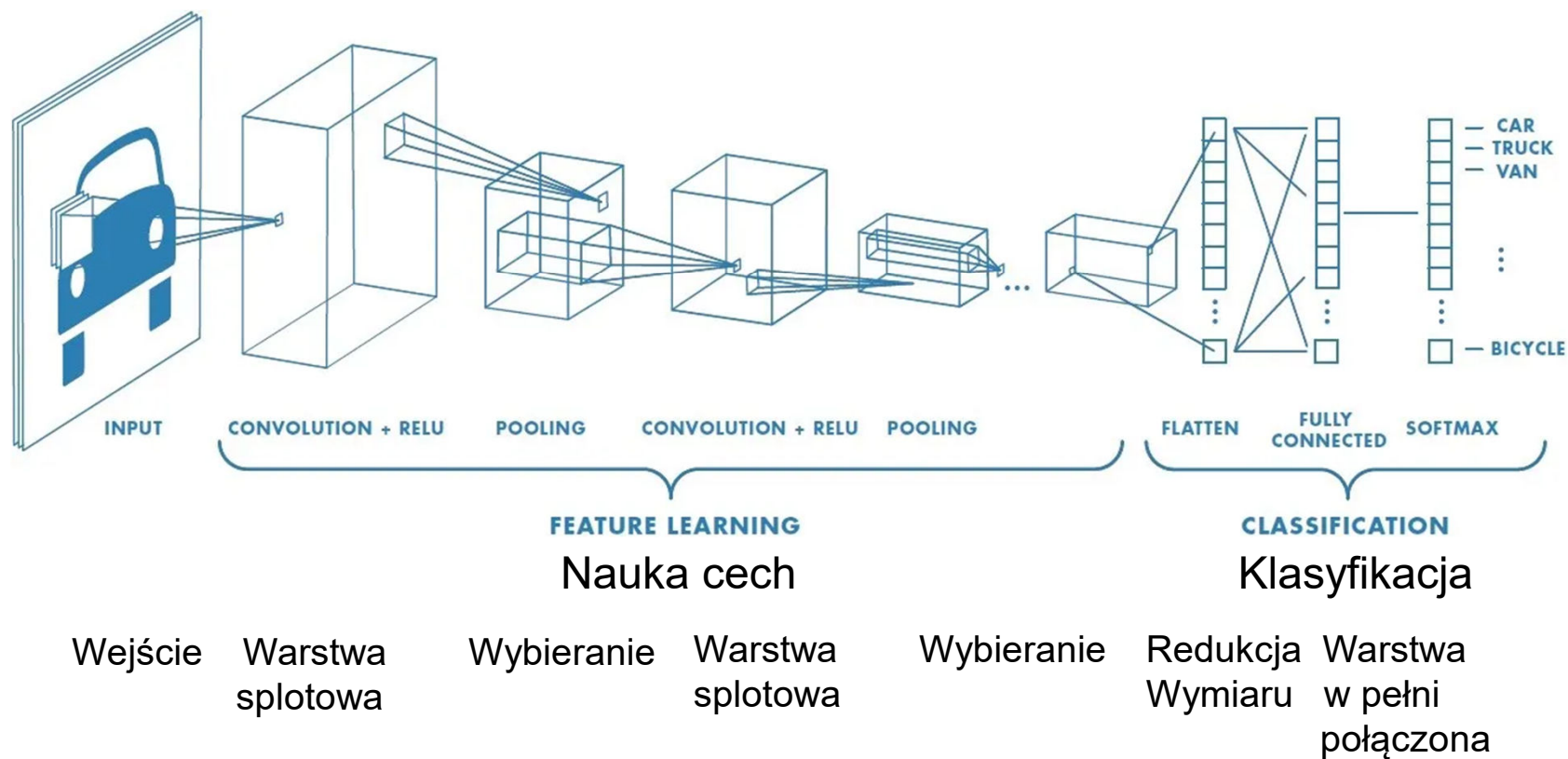
- Na wejściu: obraz.
- Na wyjściu:
 - etykieta klasy dominującego obiektu w obrazie;
 - ewentualnie prawdopodobieństwo poprawności określenia klasy.

Wykrywanie obiektu – *Object detection*

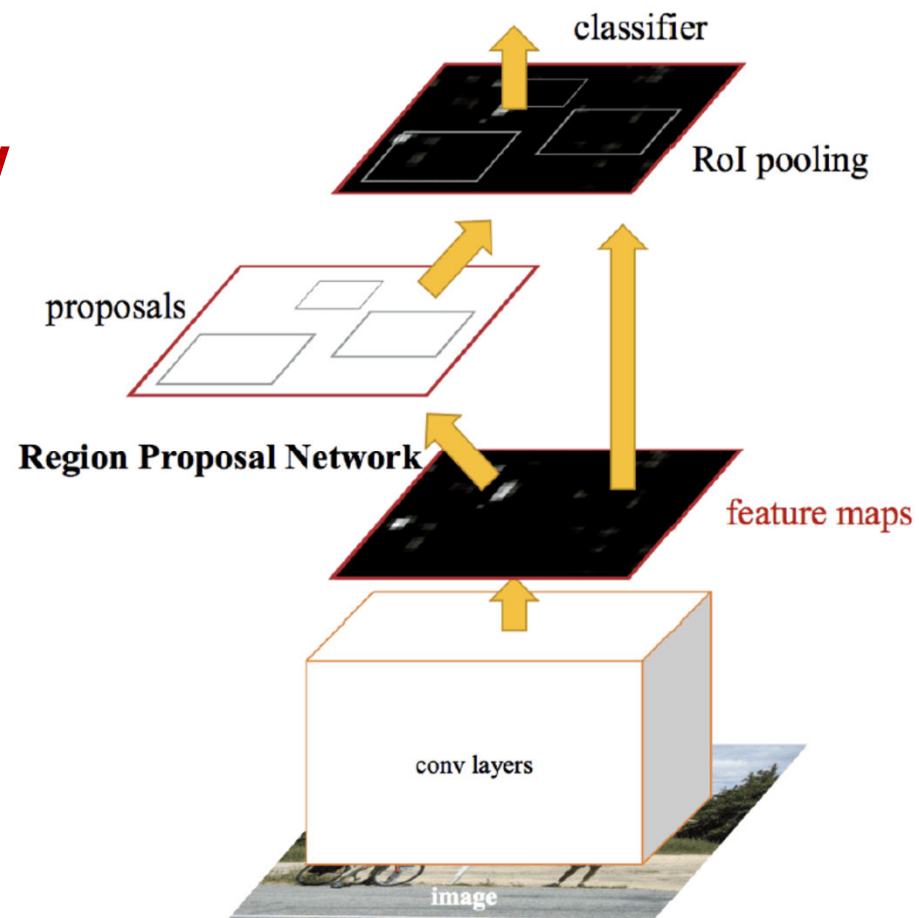
Należy odróżniać wykrywanie obiektów (*object detection*) od klasyfikacji obrazu (*image classification*).



Klasyfikacja



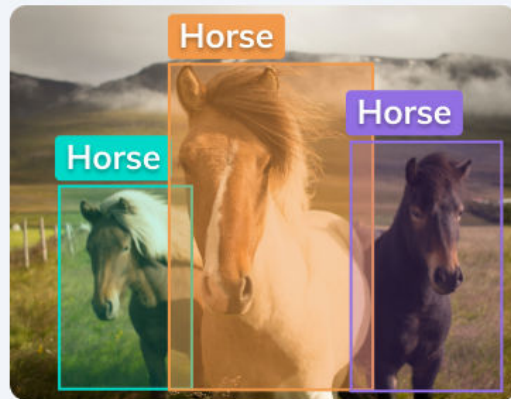
Praktyczna realizacja wykrywania obiektów



<https://machinelearningmastery.com/wp-content/uploads/2019/03/Summary-of-the-Faster-R-CNN-Model-Architecture.png>

Segmentacija instancji

**Object Detection + Semantic Segmentation
= Instance Segmentation**



Object detection



Semantic Segmentation



Instance Segmentation

Marek Domański

Materiały pomocnicze do wykładu **Wprowadzenie do multimedków** **Politechnika Poznańska**

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimedków* prowadzonych przez prof. Marka Domańskiego na Wydziale Informatyki i Telekomunikacji.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

Badania w zakresie wizji w Instytucie Telekomunikacji Multimedialnej



dr hab. inż. Dawid
Mieloch,
profesor PP
Dyrektor Instytutu



dr inż. Sławomir
Maćkowiak
**Z-ca Dyrektora
Instytutu**



prof. dr hab. inż.
Marek Domański

Badania w zakresie wizji oraz implementacji w układach elektronicznych



dr hab. inż.
Mieczysław Jessa, prof. PP



dr hab. inż.
Olgierd Stankiewicz



dr hab. inż
Adrian Dziembowski, prof. PP



dr inż.
Krzysztof Klimaszewski



dr inż.
Tomasz Grajek



dr hab. inż
Damian Karwowski

Zespół Multimediiw

Tematyka badań:

- Obraz i wizja – przetwarzanie, kompresja.
- Analiza obrazów przestrzennych – estymacja głębi, przetwarzanie wstępne obrazów wielowidokowych.
- Przetwarzanie i kompresja wizji wielowidokowej i trójwymiarowej.
- Wirtualna rzeczywistość – synteza widoków wirtualnych, przesyłanie obrazów wszechkierunkowych, estymacja głębi w systemach wszechkierunkowych, przetwarzanie obrazów wszechkierunkowych.
- Projektowanie systemów z układami programowalnymi – zwłaszcza dla multimediiów.
- Wykorzystanie metod sztucznej inteligencji.
- Kodowanie wizji dla maszyn (IoT).
- Zastosowanie metod głębokiego uczenia w kompresji wizji.
- Analiza wizji dla systemów dozoru wizyjnego.

Zespół Telekomunikacji Multimedialnej

Współpraca z przemysłem

Partnerzy :

- ADB – Advanced Digital Broadcast, Polska
- Halliburton USA,
- Mitsubishi Electric (Wlk. Brytania / Japonia),
- Samsung Korea,
- TVP Polska,
- IMITPI Polska, 3D Vision Polska,
- Huawei (Niemcy / Chiny),
- OPPO Chiny,
- Fraunhofer Institute HHI – Niemcy,
- Electronics and Telecommunications Research Institute (Korea).

Serwerownia

- 10 serwerów plików – przestrzeń ponad 300TB,
- 3 serwery odpowiedzialne za usługi dostępu do serwerów plików i usługi zewnętrzne,
- Urządzenia i przełączniki pracujące z przepływnościami interfejsów 10Gbps (światłowodowe oraz kablowe),
- Dedykowana instalacja kablowa do pomieszczeń w standardzie 6 oraz 7 kategorii (prędkość 10 Gbps),
- Ponad 30 jednostek obliczeniowych (częściowo również z wydajnymi kartami graficznymi),
- Wydajne chłodzenie pomieszczenia.

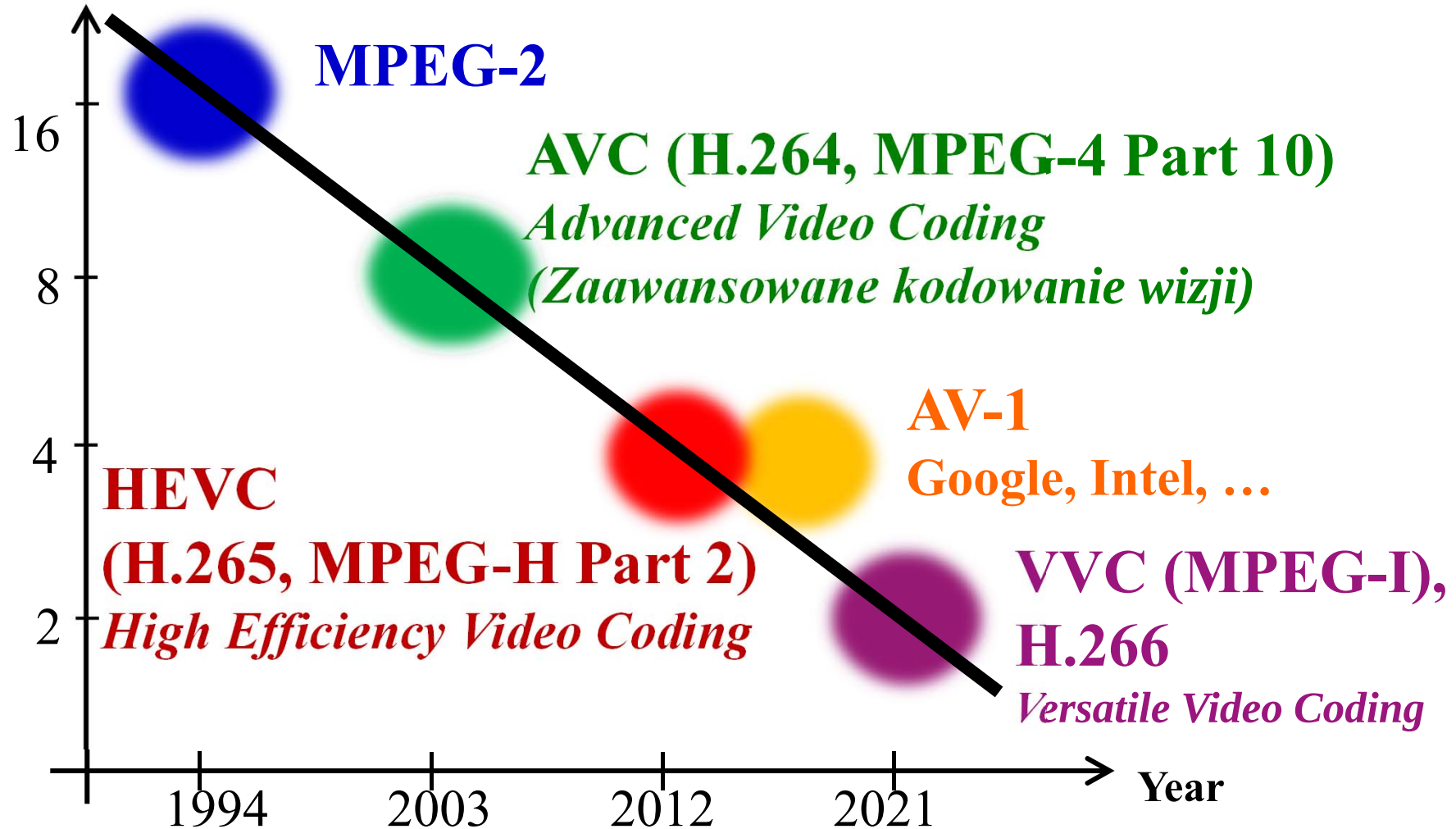


Przygotowanie norm międzynarodowych ISO/IEC/ITU

- Udział w konkursach.
- Przygotowanie norm.
- Projekty badawcze (udział studentów i doktorantów).
- W ostatnich 20 latach byliśmy w zakresie przygotowania technik dla normalizacji najaktywniejszym ośrodkiem naukowym na obszarze pomiędzy Niemcami a Chinami.

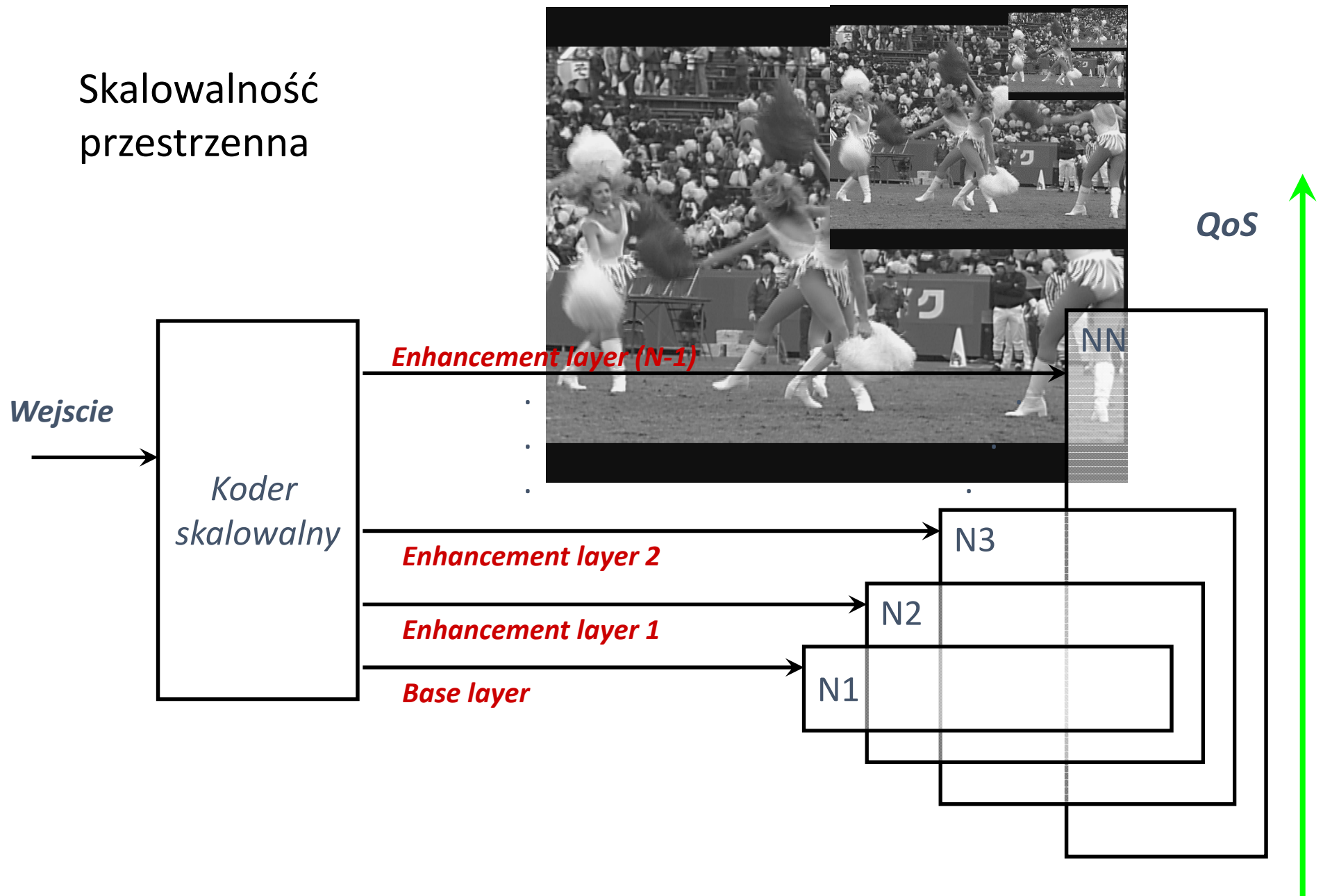
Generacje technik kompresji

Przybl. szybkość bitowa dla **HD** (1920×1080, 25 frame/s) [Mb/s]



KODOWANIE SKALOWALNE ZAGNIEŹDŻANIE STRUMIENI

Skalowalność
przestrzenna

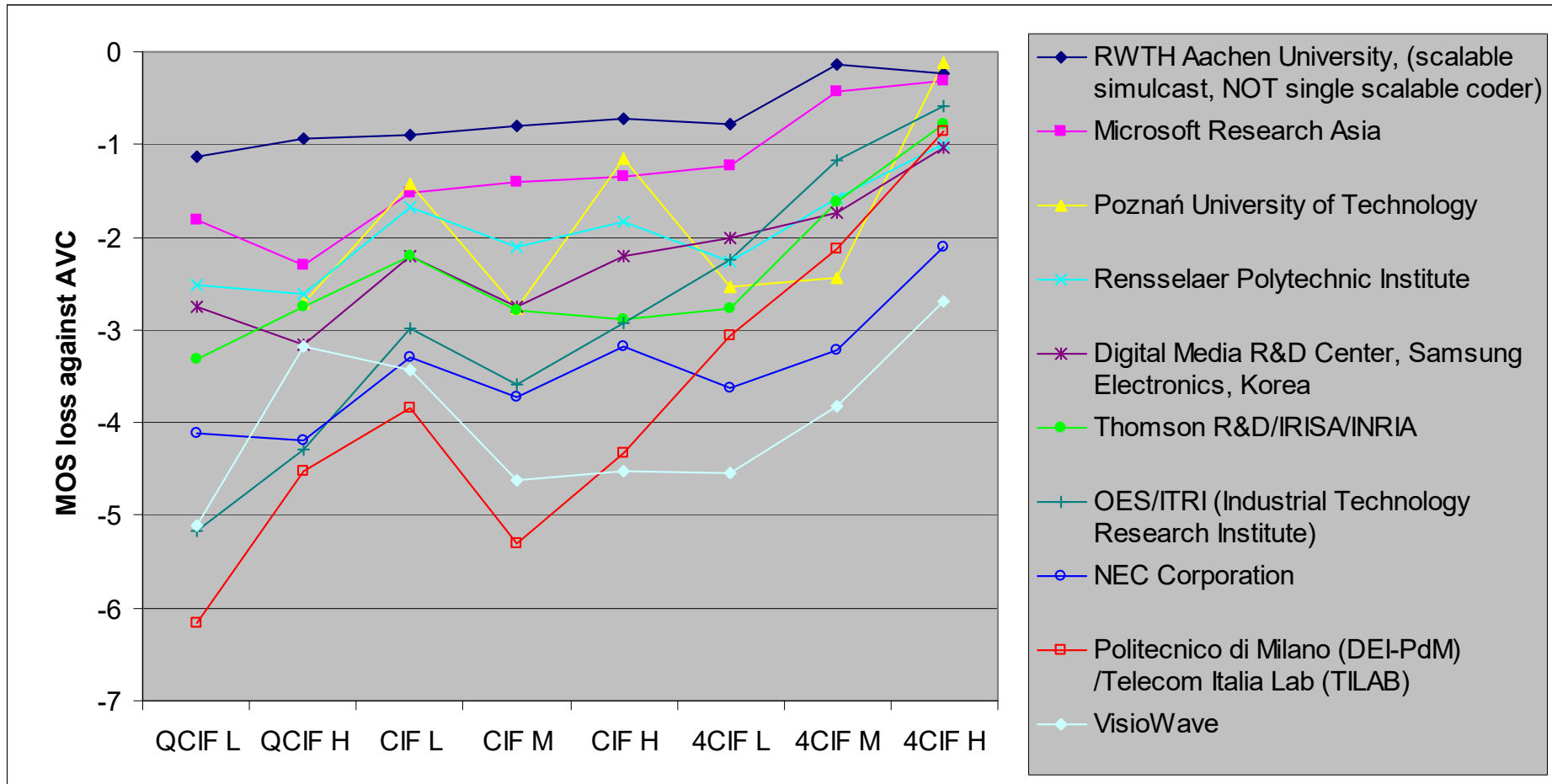


Monachium 2004

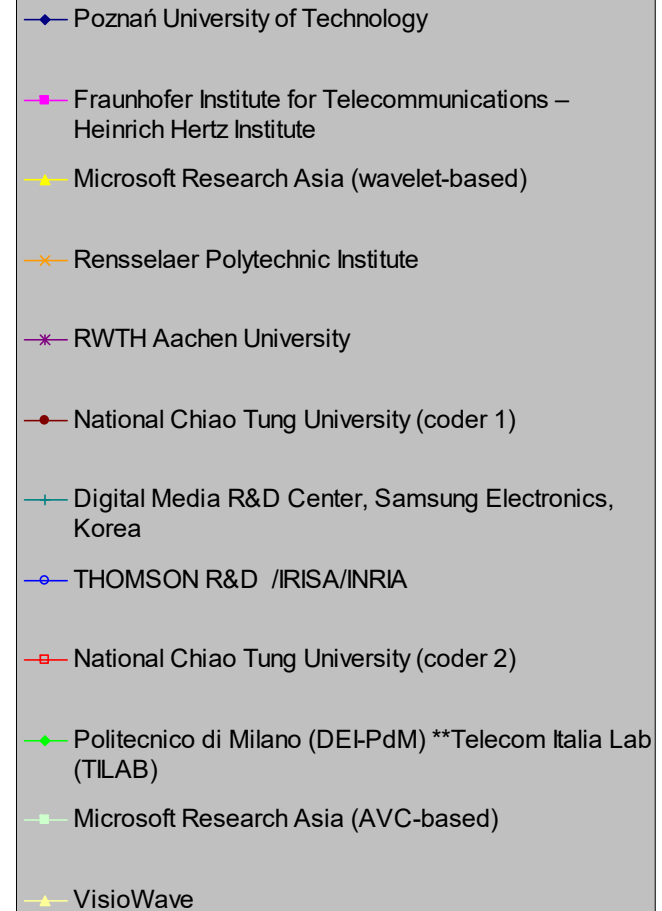
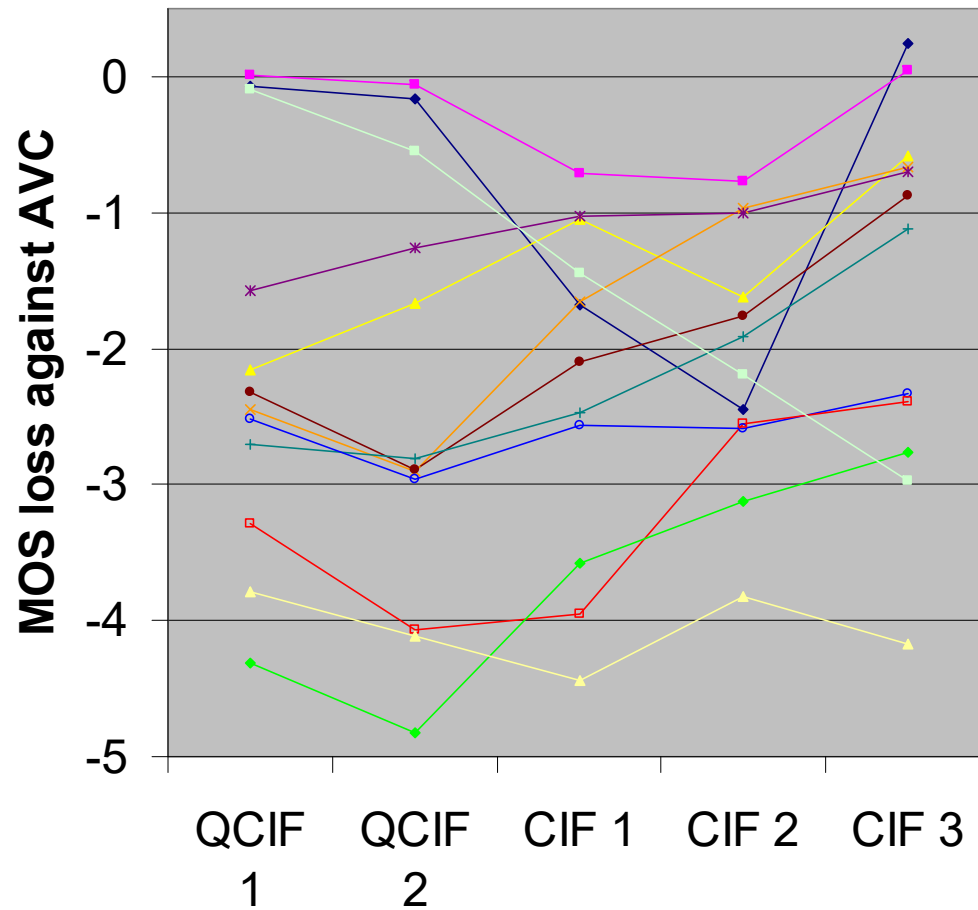
Propozycje w 2 grupach:

- Skalowalne rozszerzenia techniki AVC.
- Niekompatybilne ze standardami kompresji techniki wykorzystujące kodowanie falkowe.

Wyniki testów dla scenariusza 1

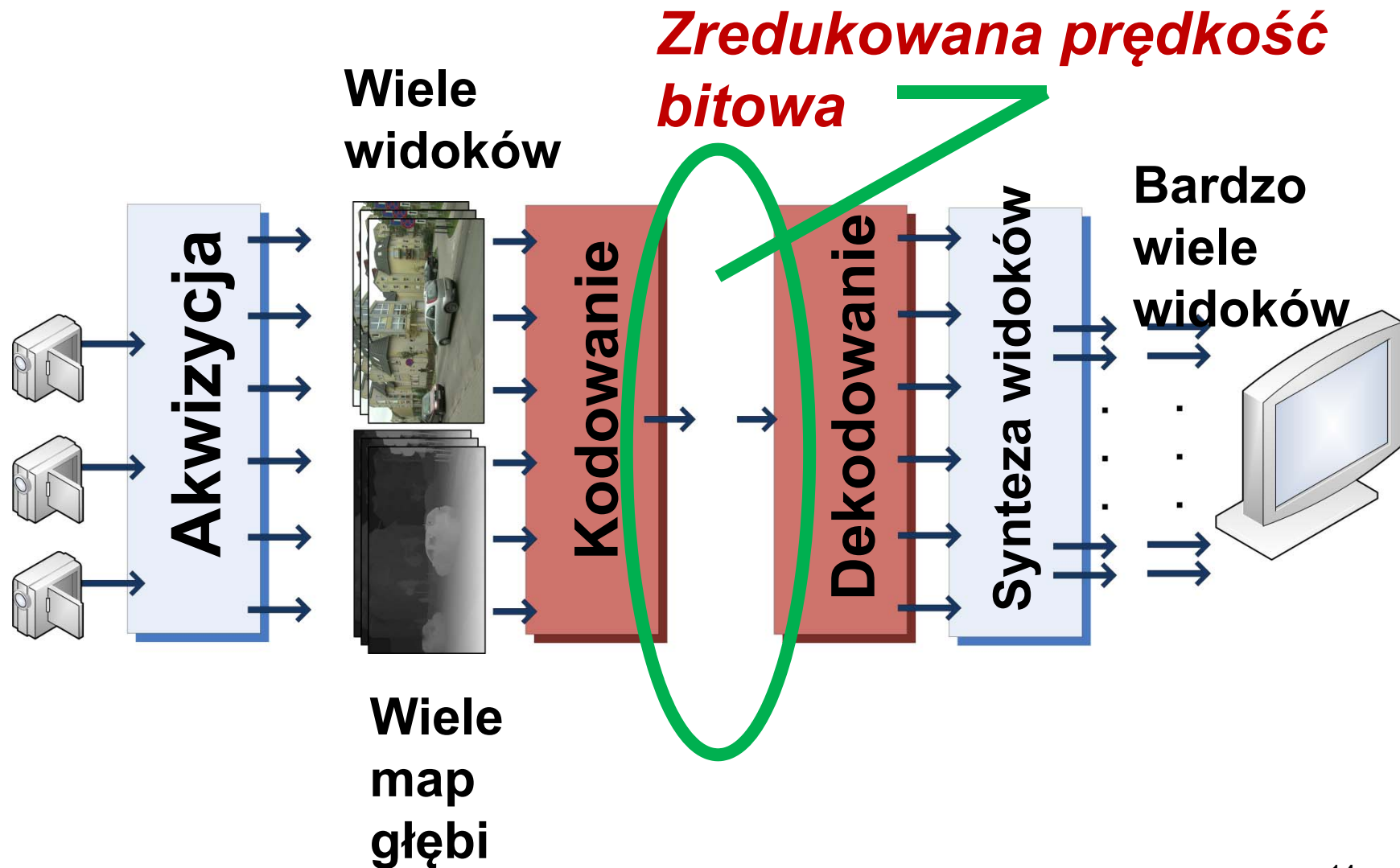


Wyniki testów dla scenariusza 2



Kodowanie sekwencji wielowidokowych i trójwymiarowych

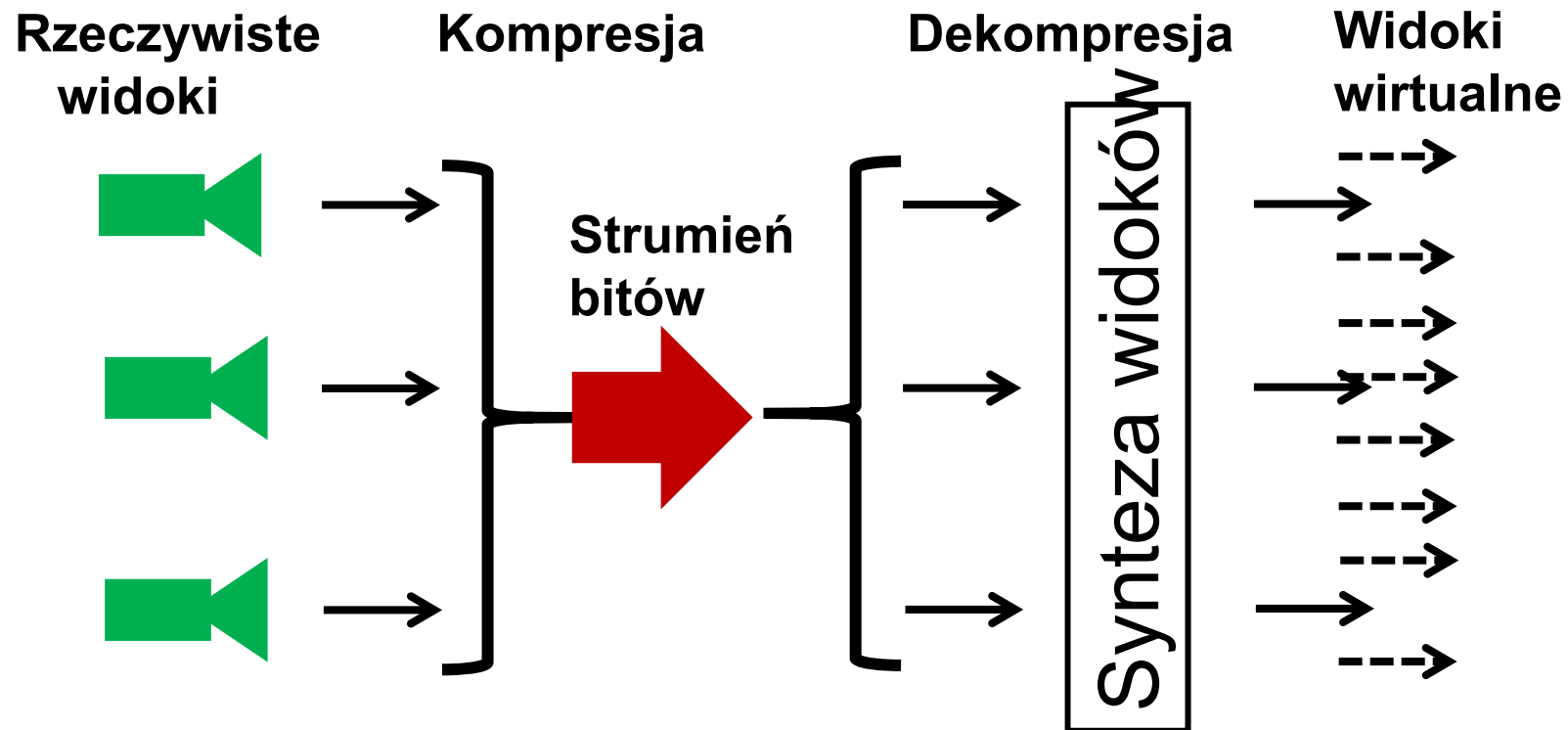
Przyszłe systemy



Problem

- **Kompresja (*multiview + depth* = MVD)**

Synteza widoków w odbiorniku:



Widok 1

Widok 2

Obrazy (tekstura) :



Głębia :



Przygotowania normy kodowania obrazu trójwymiarowego

- **Na bazie technologii HEVC.**
- **Grupa MPEG ogłosiła wezwanie do zgłaszania technik kompresji.**
- **W listopadzie 2011: wyniki porównań.**
- **Najlepsze techniki z dwóch ośrodków:
Inst. Fraunhofera HHI i Politechnika Poznańska.**

Badanie propozycji

- Każda propozycja:
Dekoder, sekwencje zdekodowane.
- Badania:
 - badania jakości **subiektywnej**,
 - dotychczas największy eksperyment
pomiaru jakości obrazu stereoskopowego.

Proposal testing

- **Each proponent:**

For 8 sequences and 4 bitrates: bitstreams,
decoded video, decoder (before Aug. 31, 2011).

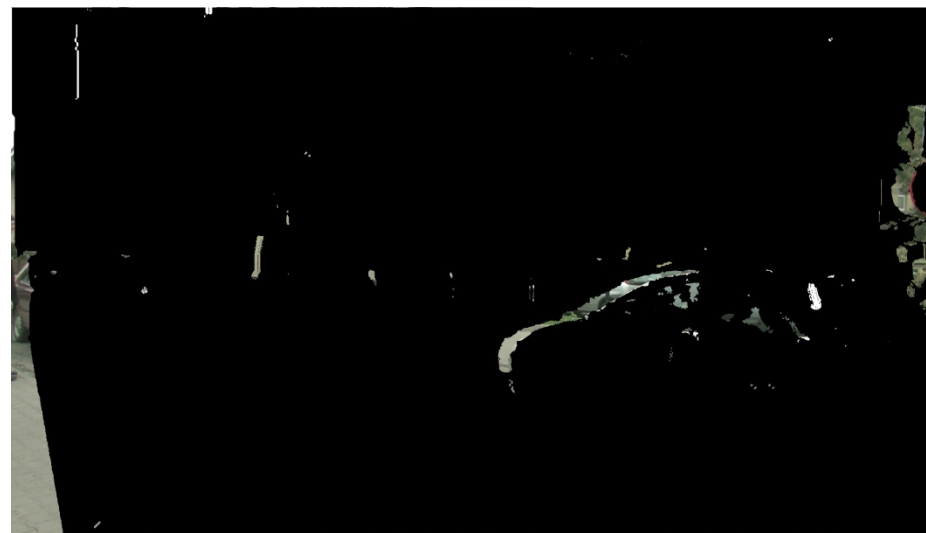
- **Testing:**

- subjective quality testing
of selected stereo pairs,
- 13 labs around the world,
- undisclosed pairing of proposals and labs .

Kodowanie obszarów odsłoniętych



Oryginalny widok boczny



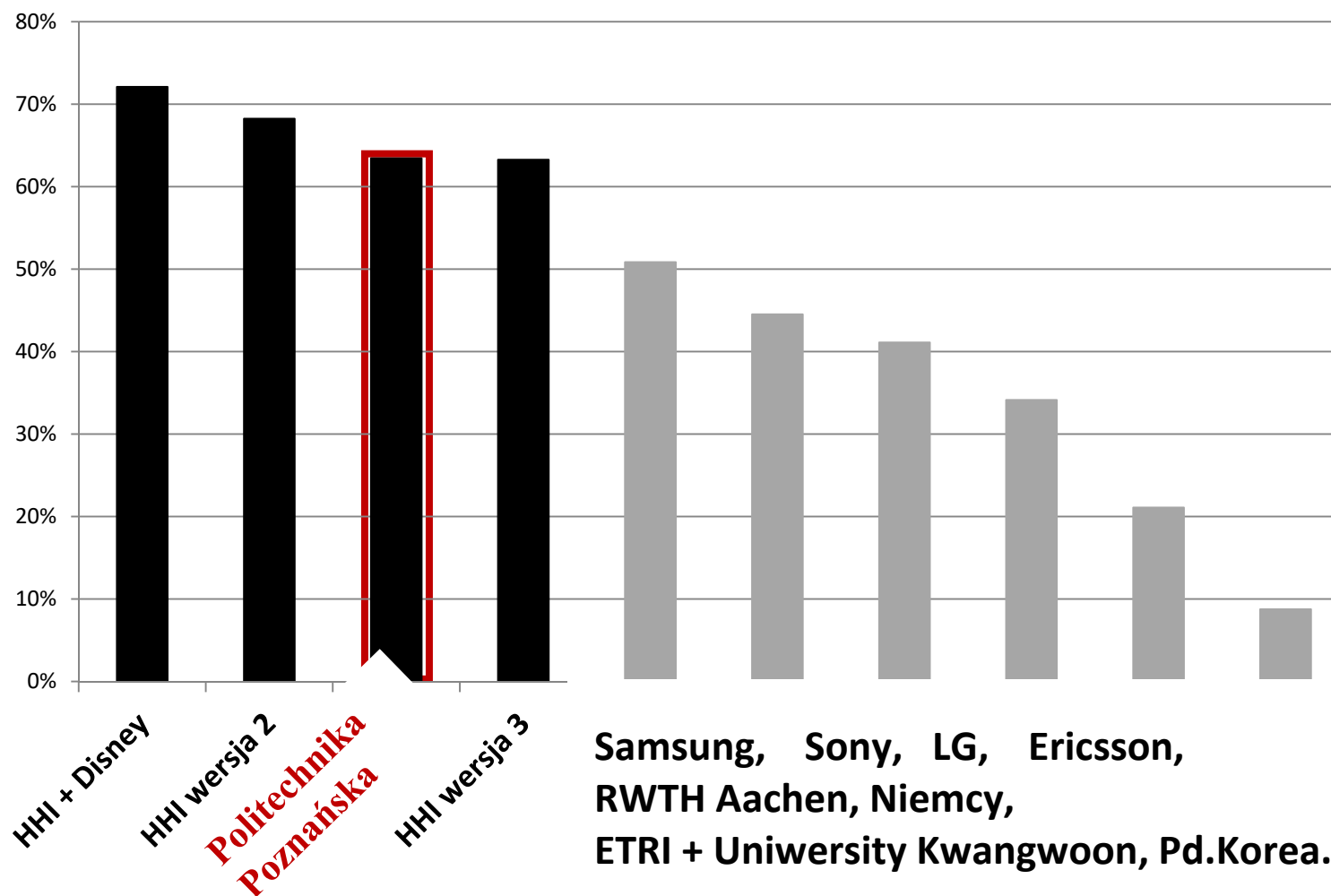
Obszary odsłonięte w bocznym widoku



Kodowane jednostki kodowania (CU)²⁰

Redukcja prędkości bitowej względem niezależnego kodowania techniką HEVC

Średnia redukcja prędkości bitowej [%]



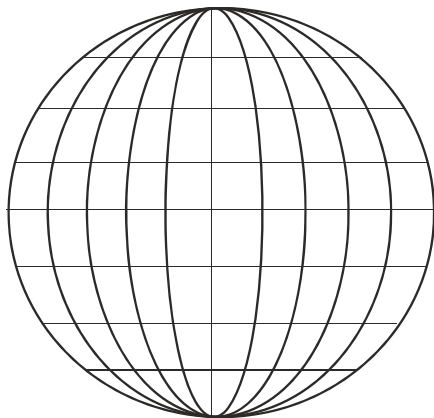
Wizja wszechogarniająca ***Immersive video***

2019

Omnidirectional cameras



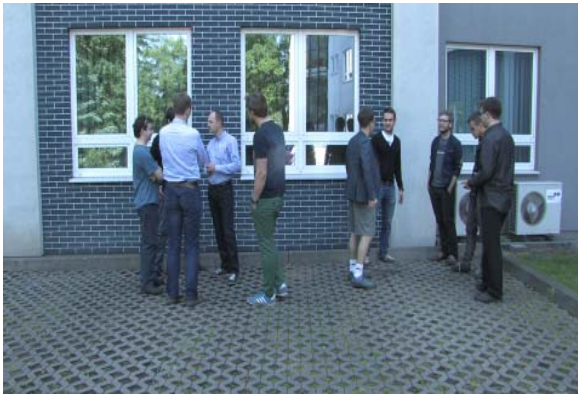
Omnidirectional video



G. Bang, G. S. Lee, N. Ho H., "Test materials for 360 3D video application discussion", *ISO/IEC JTC1/SC29/WG11 MPEG2016/M37810*, San Diego, USA, February 2016.

Multiview representation of immersive video

Input views:



View 0



View 1



View 2

From input views the output video is synthesized (or rendered).

Virtual navigation

- Viewer may watch the scene from arbitrary position and arbitrary direction (both are changeable over time).
- For a given frame of video: the view direction and virtual viewer location form a **viewport**.



Artifacts visible: High quality virtual navigation is still challenging.

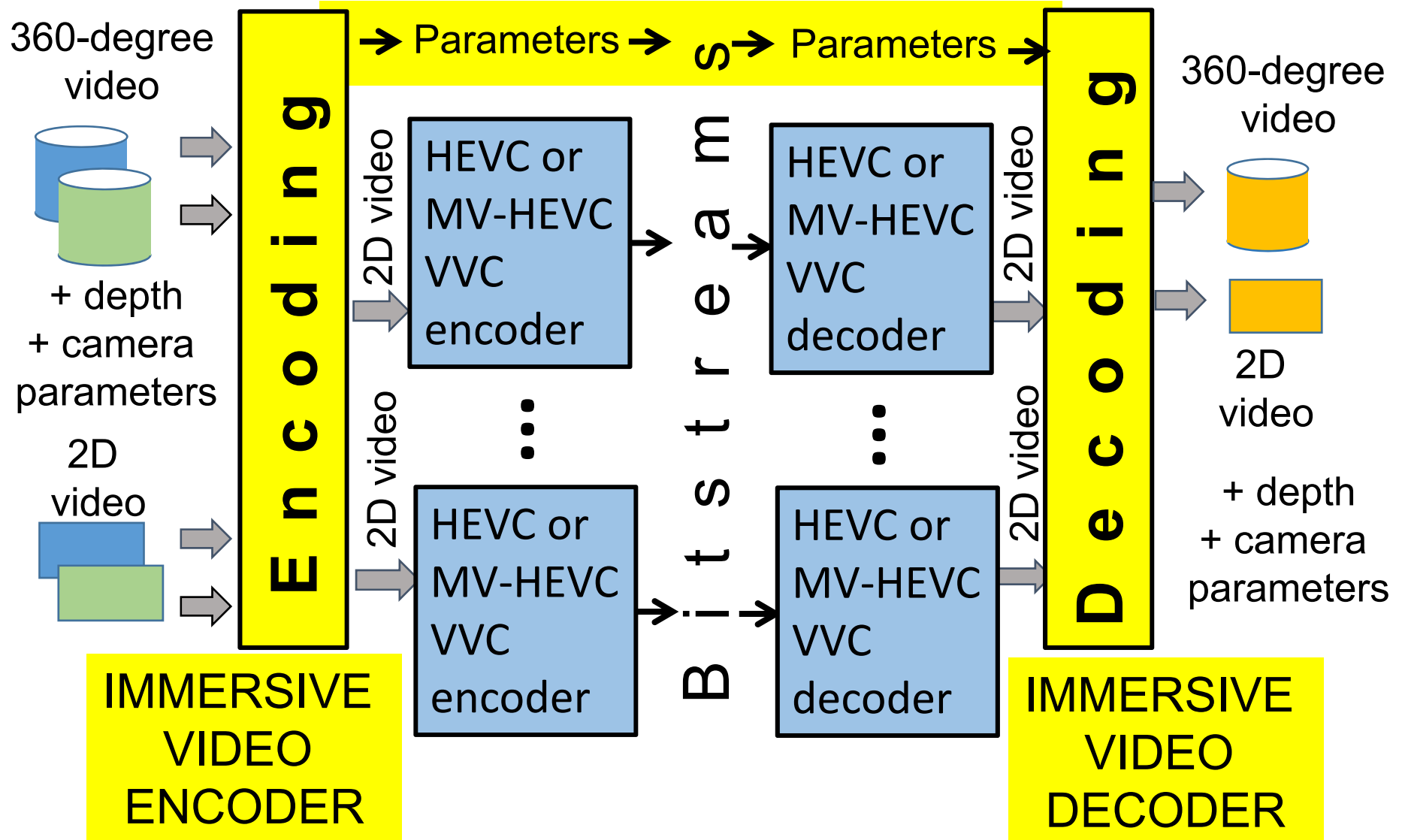
Virtual navigation

- Viewer may watch the scene from arbitrary position and arbitrary direction (both are changeable over time).
- For a given frame of video: the view direction and virtual viewer location form a **viewport**.



Artifacts visible: High quality virtual navigation is still challenging.

Immersive video compression





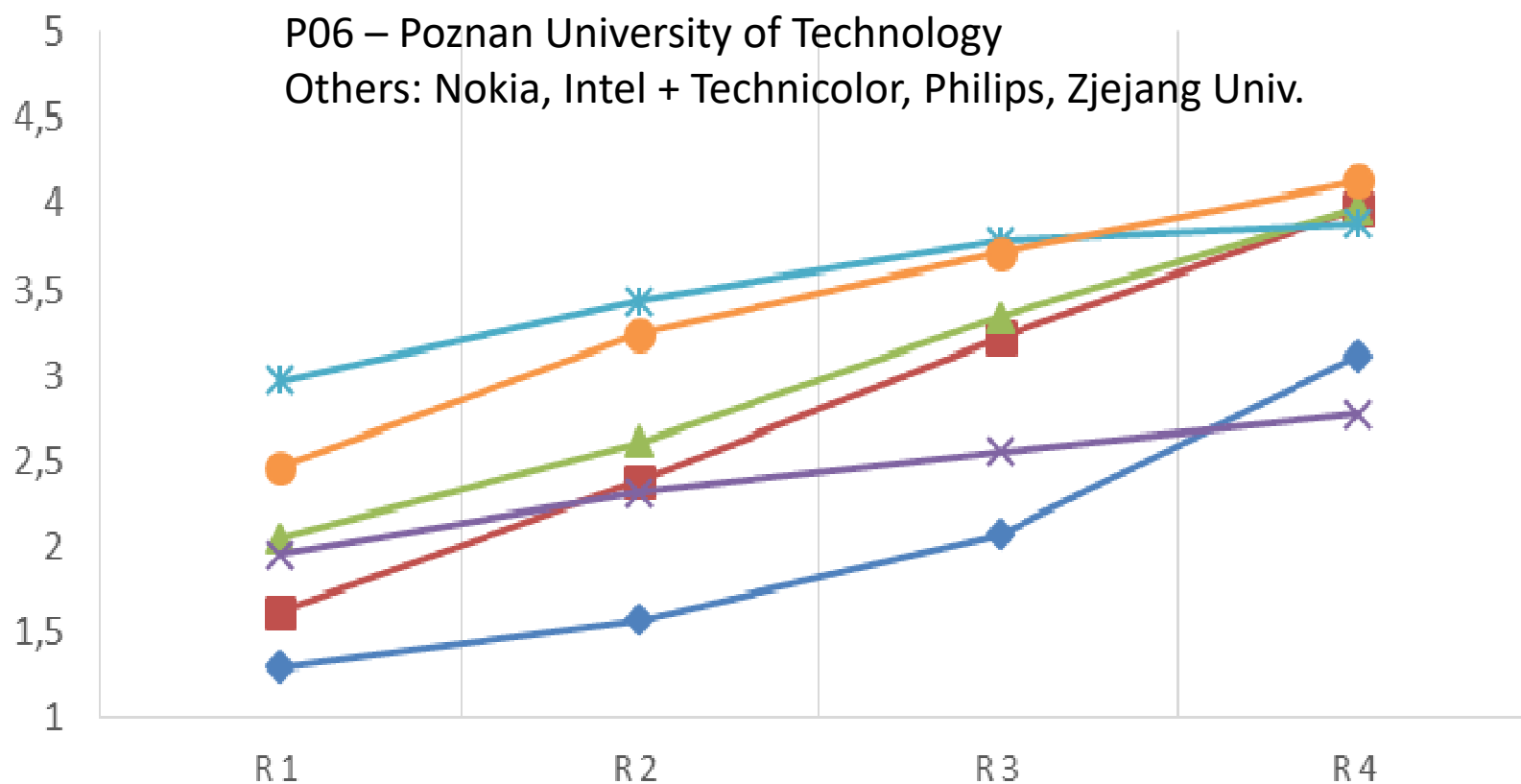


ROMA (NAIVE): AVERAGE MOS

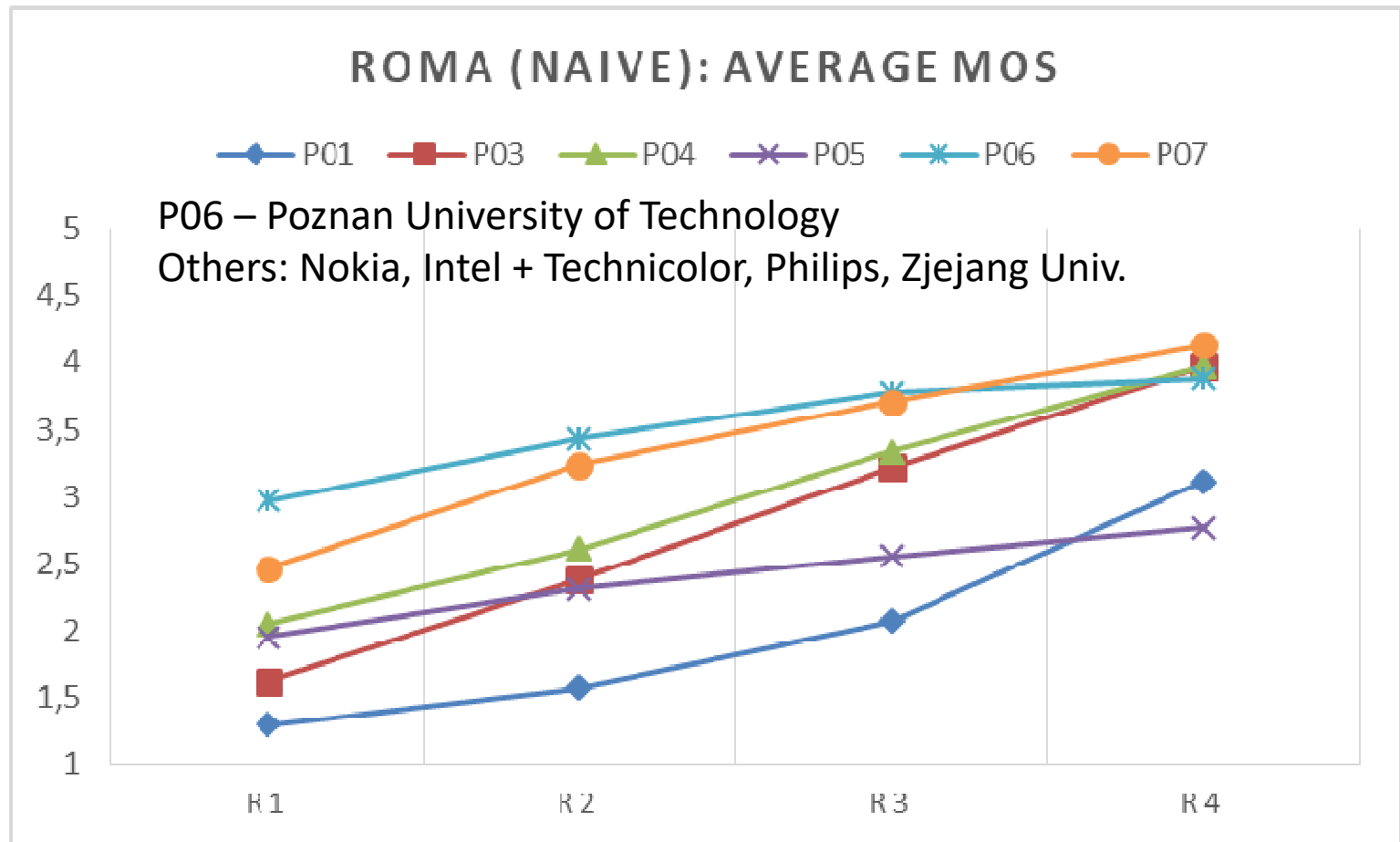
—◆— P01 —■— P03 —▲— P04 —×— P05 —*— P06 —●— P07

P06 – Poznan University of Technology

Others: Nokia, Intel + Technicolor, Philips, Zhejiang Univ.



Plot 2 provides rank values: A proponent gets 1, if the proposal was scored with the highest MOS value, he gets 2, if the respective proposal was scored with the second highest MOS value etc. The highest the rank, the lower the MOS score.

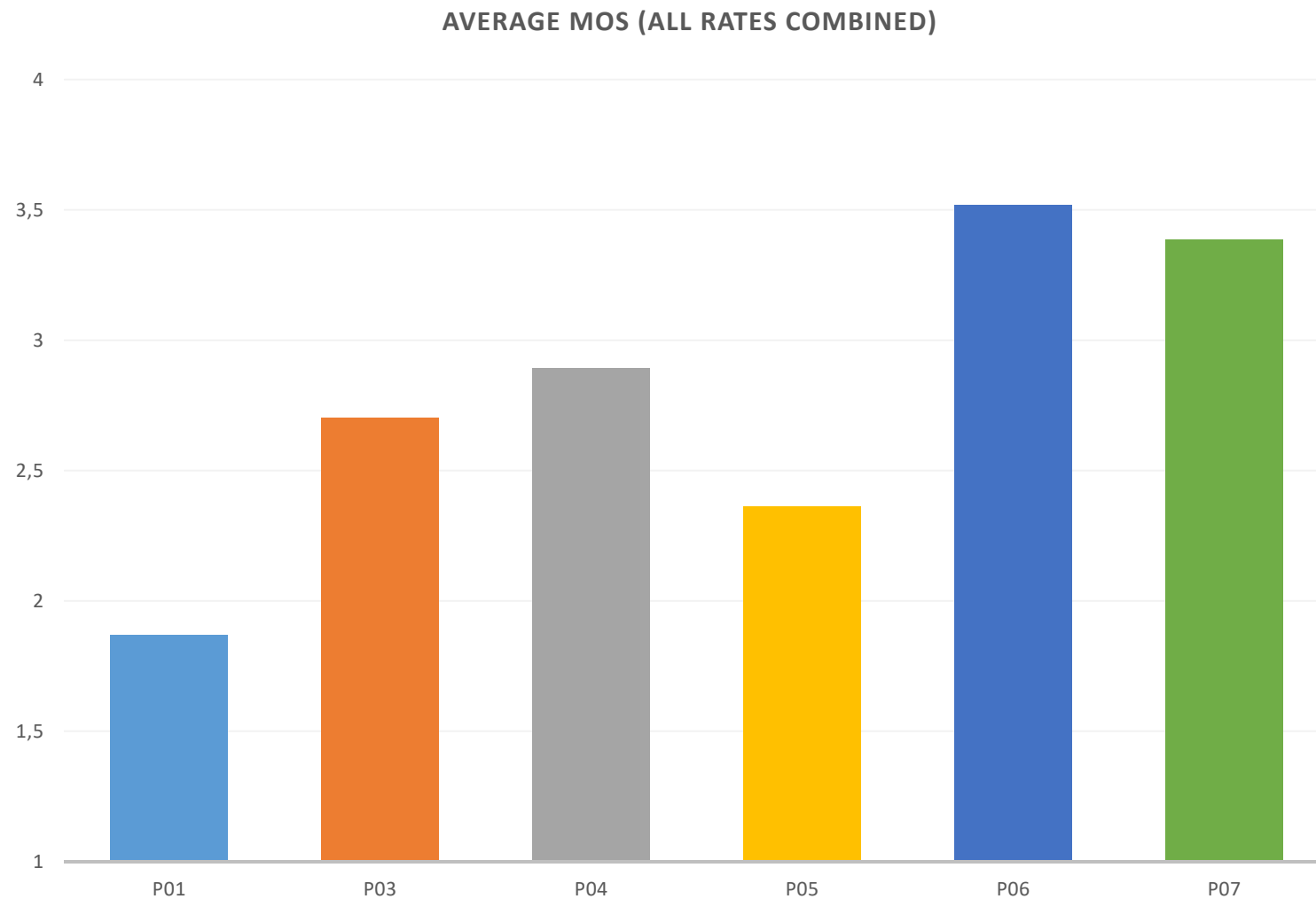




Ranking of the proposals



Average MOS values over poses, sequences and bitrates

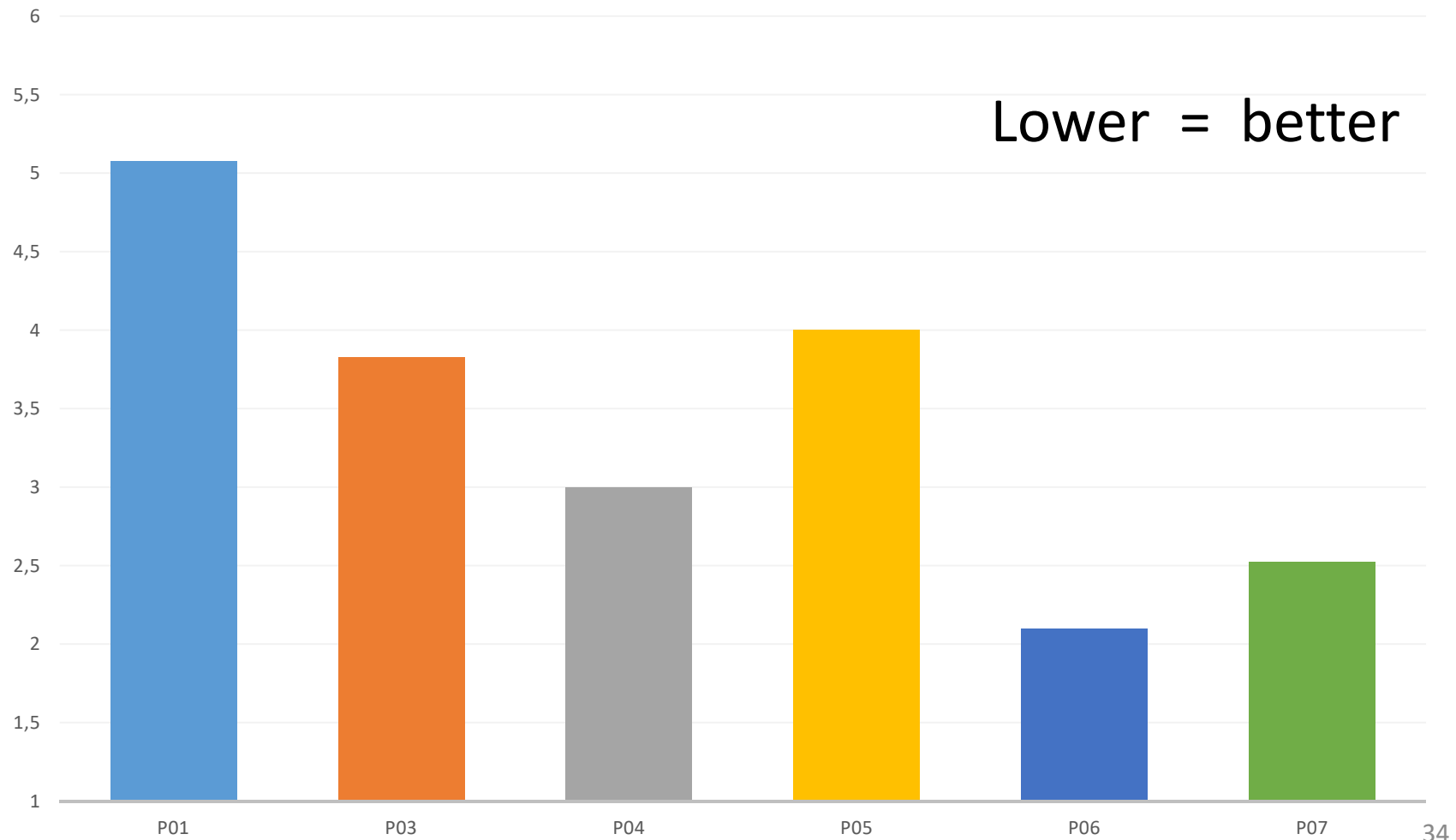




Ranking of the proposals



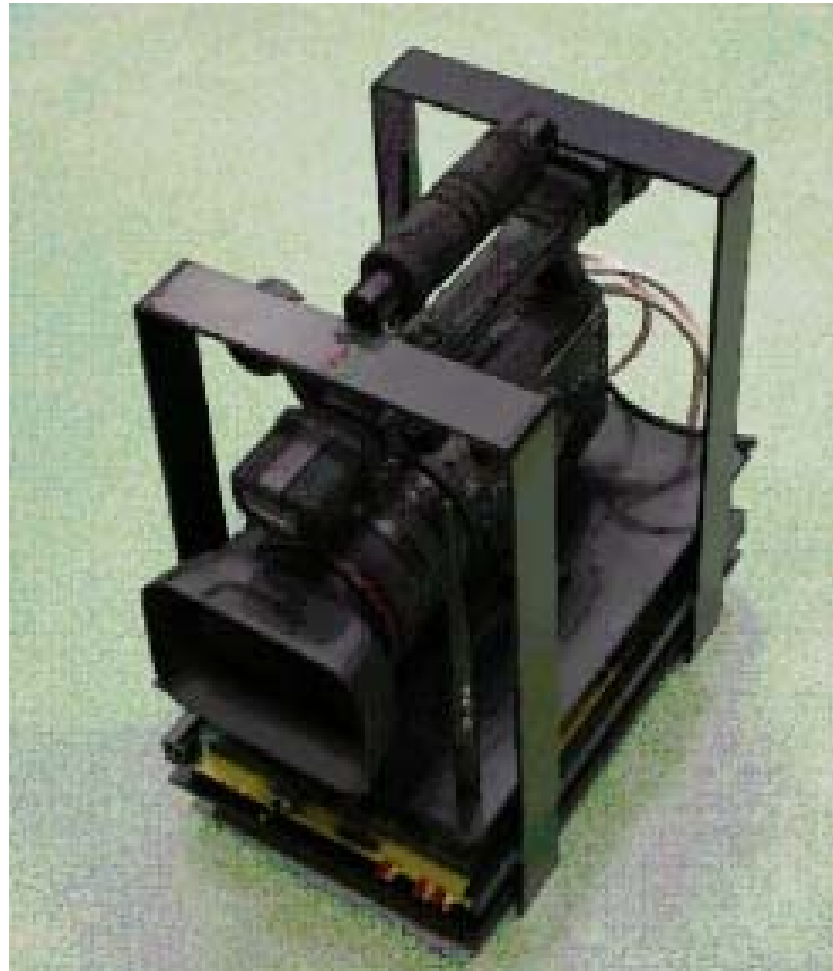
Average proponent positions over poses, sequences and bitrates



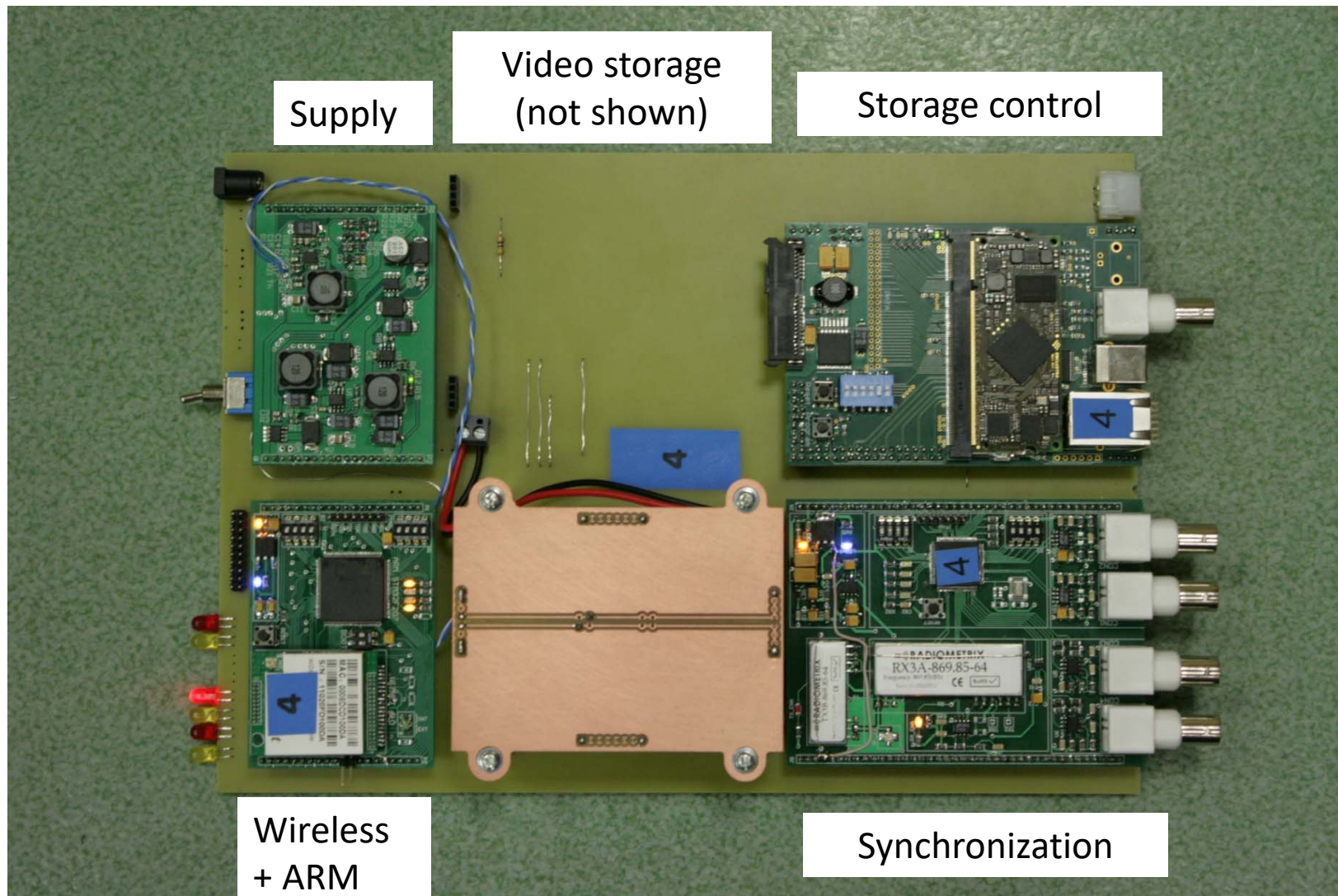
Akwizycja testowych sekwencji



Wireless camera module



Wireless camera module



Sekwencje demonstracyjne (obrazy uzyskane za pomocą zsynchronizowanych kamer)



Sekwencje demonstracyjne (obrazy uzyskane za pomocą zsynchronizowanych kamer)



Telewizja swobodnego punktu widzenia - Free-Viewpoint Television (FTV) Wirtualna nawigacja

- Scena rejestrowana zestawem kamer z mikrofonami
- Widz :
 - wirtualnie przemieszcza się w scenie
 - może oglądać scenę z DOWOLNEGO miejsca
- Każdy (wirtualny) widok jest syntezyzowany na podstawie obrazów rejestrowanych przez rzeczywiste kamery



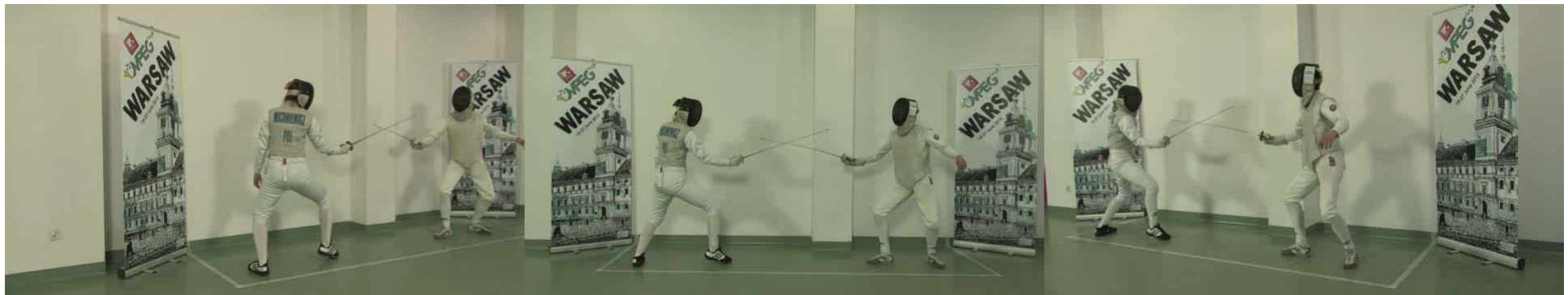
Telewizja swobodnego punktu widzenia - Free-Viewpoint Television (FTV) Wirtualna nawigacja

- Scena rejestrowana zestawem kamer z mikrofonami
- Widz :
 - wirtualnie przemieszcza się w scenie
 - może oglądać scenę z DOWOLNEGO miejsca
- Każdy (wirtualny) widok jest syntezyzowany na podstawie obrazów rejestrowanych przez rzeczywiste kamery



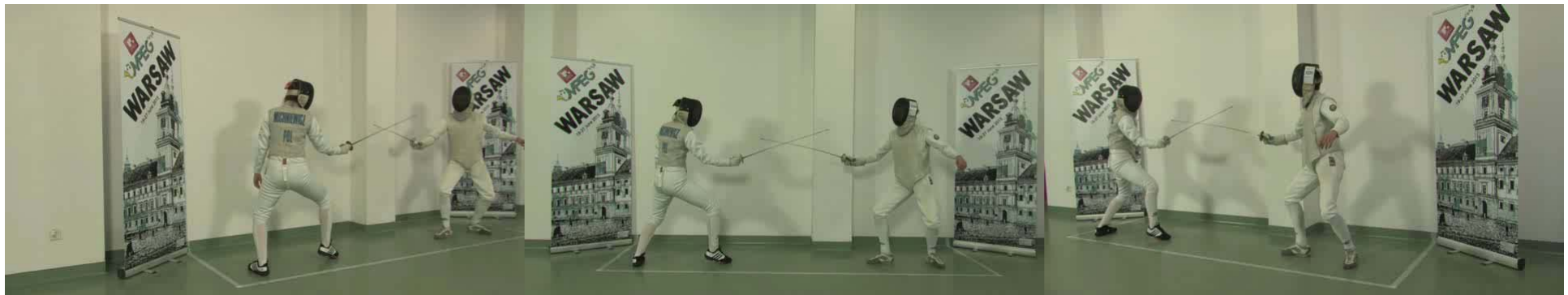
Multiview video

Selected 3 views

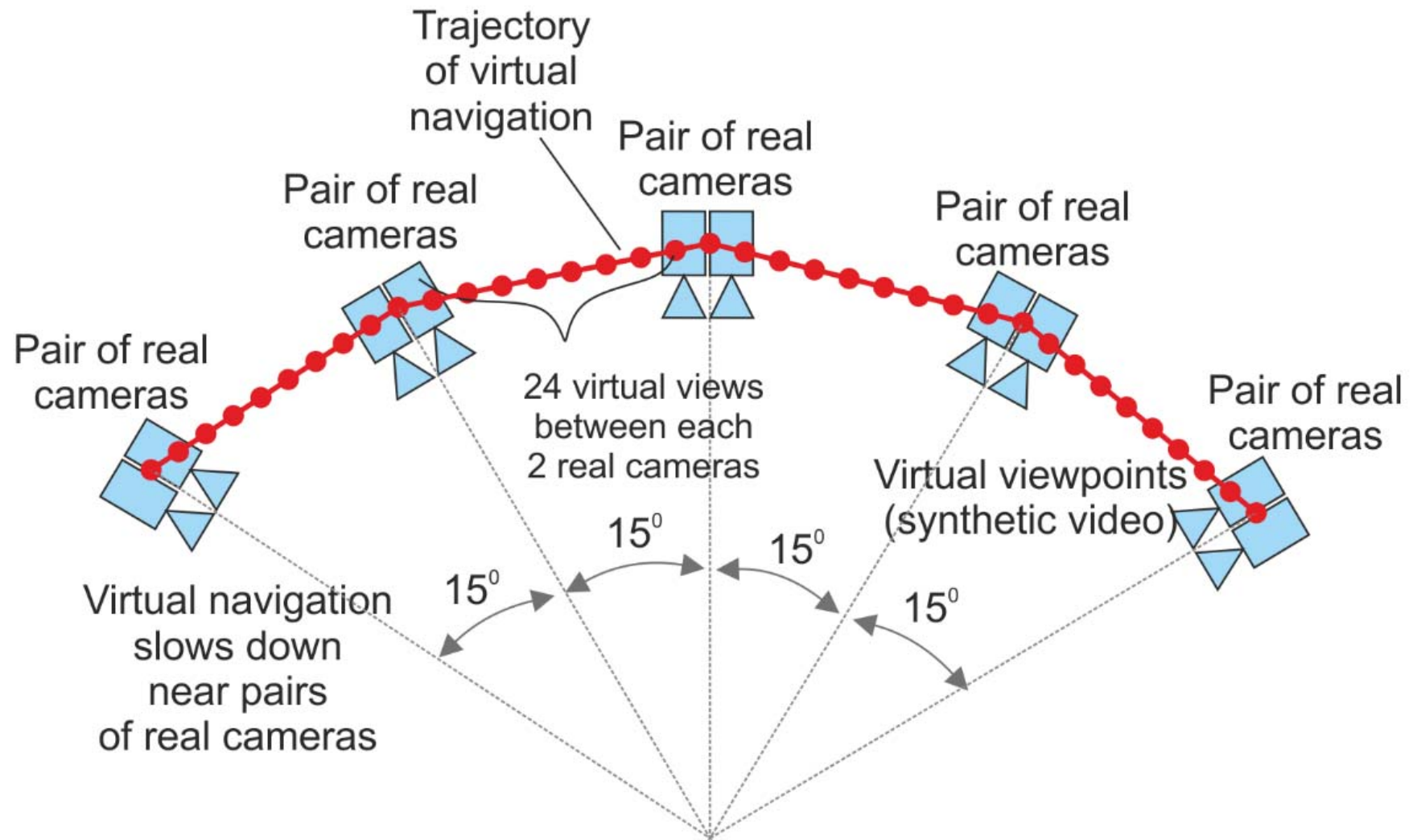


Multiview video

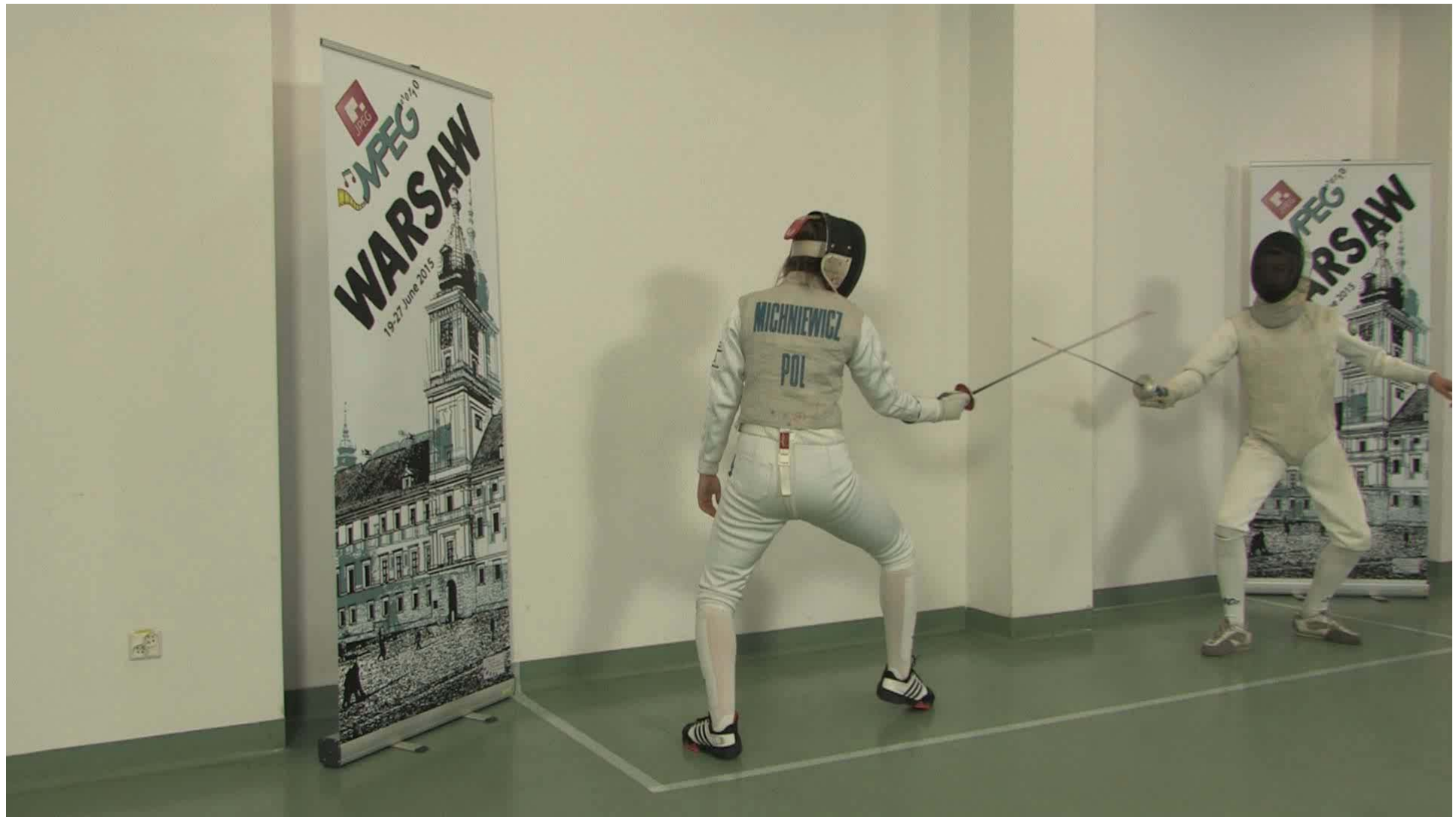
Selected 3 views



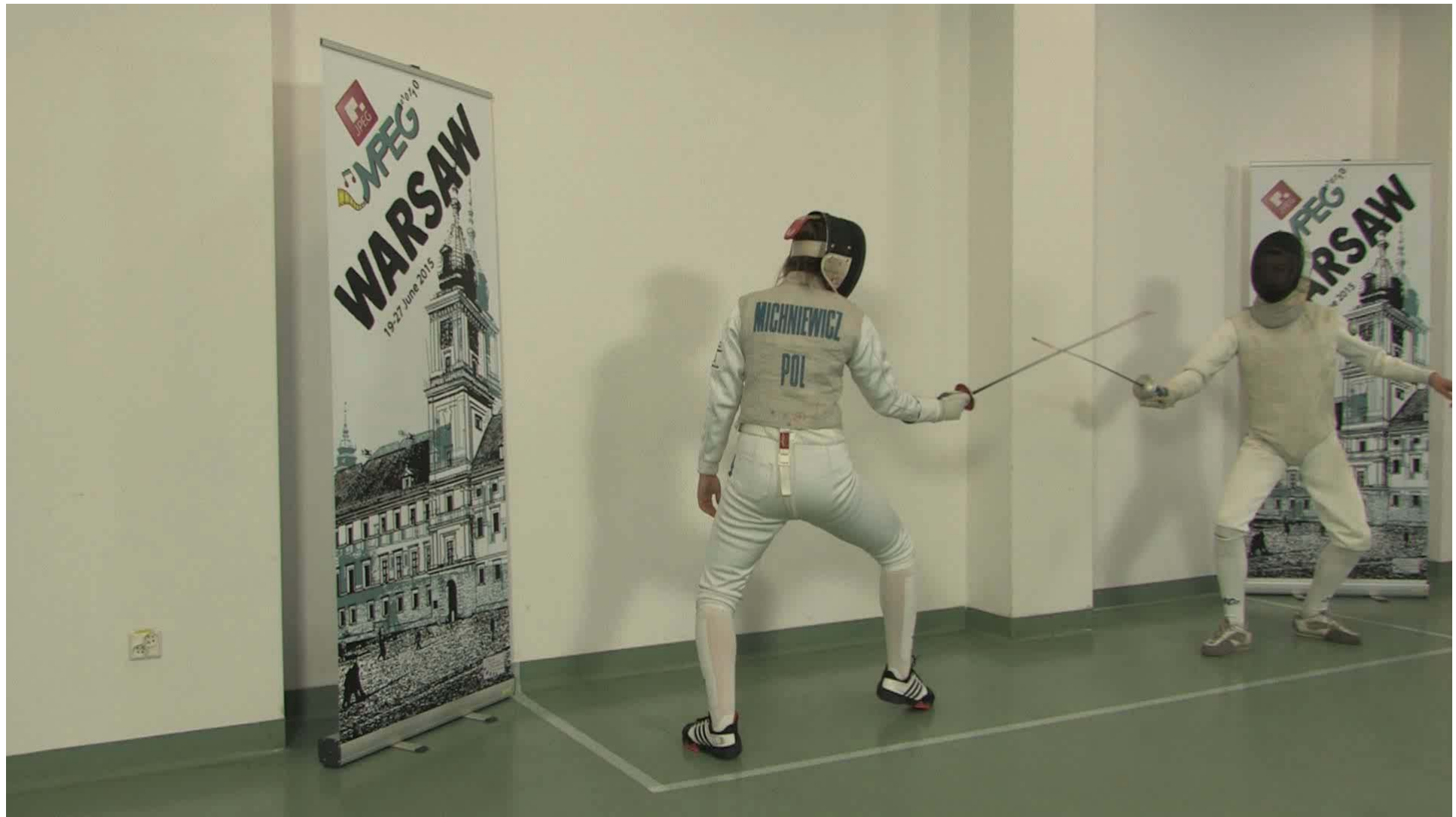
Rendered virtual views



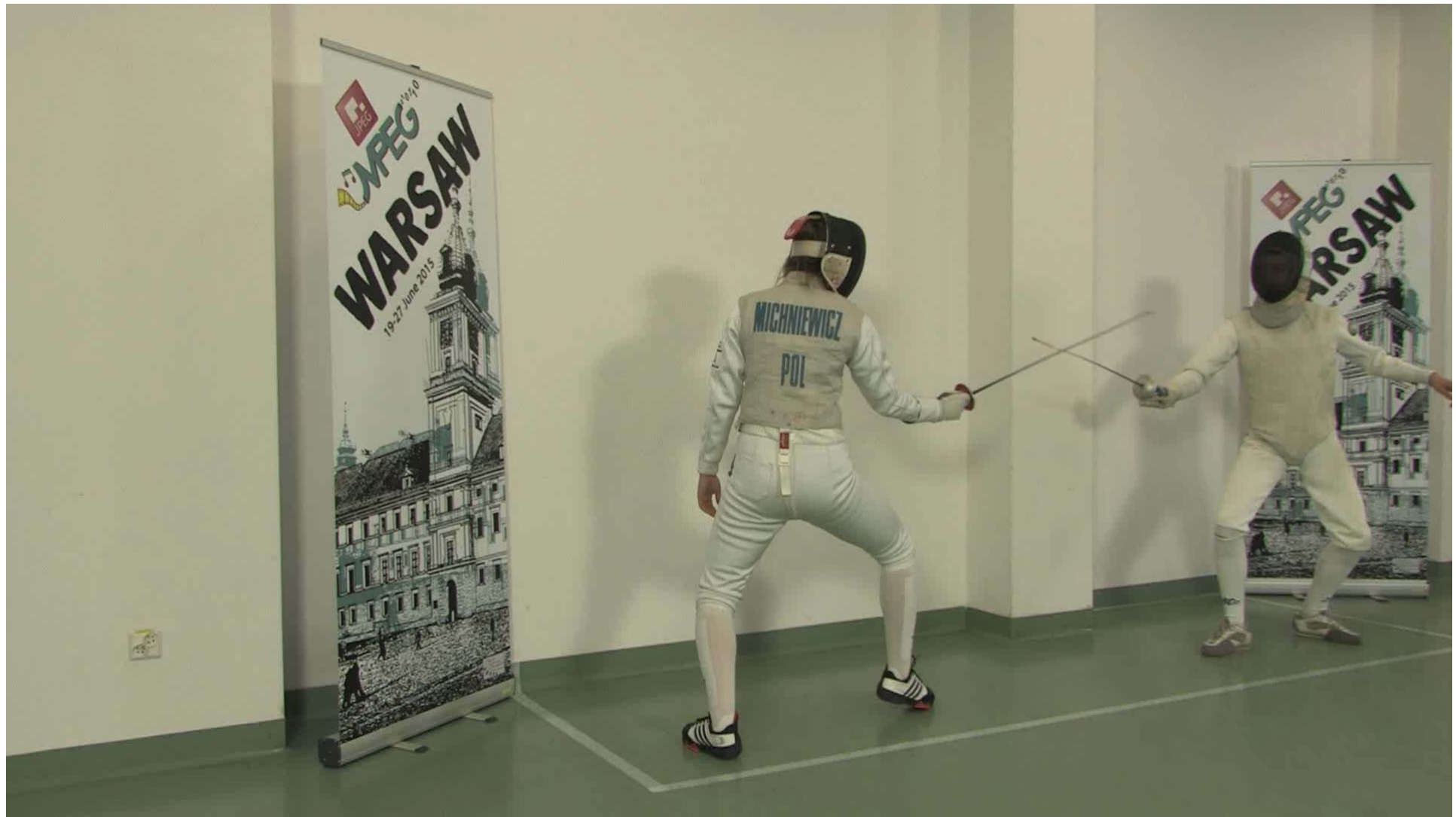
Virtual navigation for Poznan_Fencing2



Virtual navigation for Poznan_Fencing2



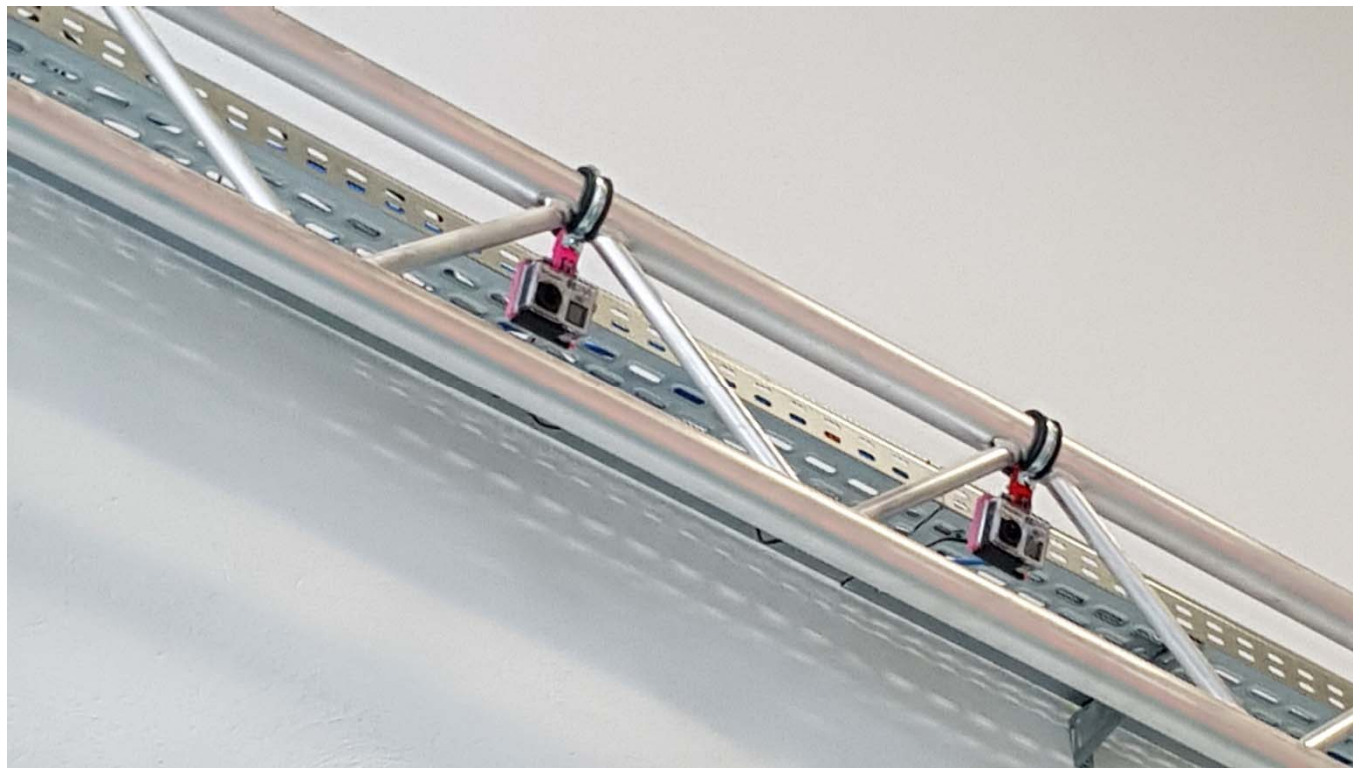
Virtual navigation for Poznan_Fencing2



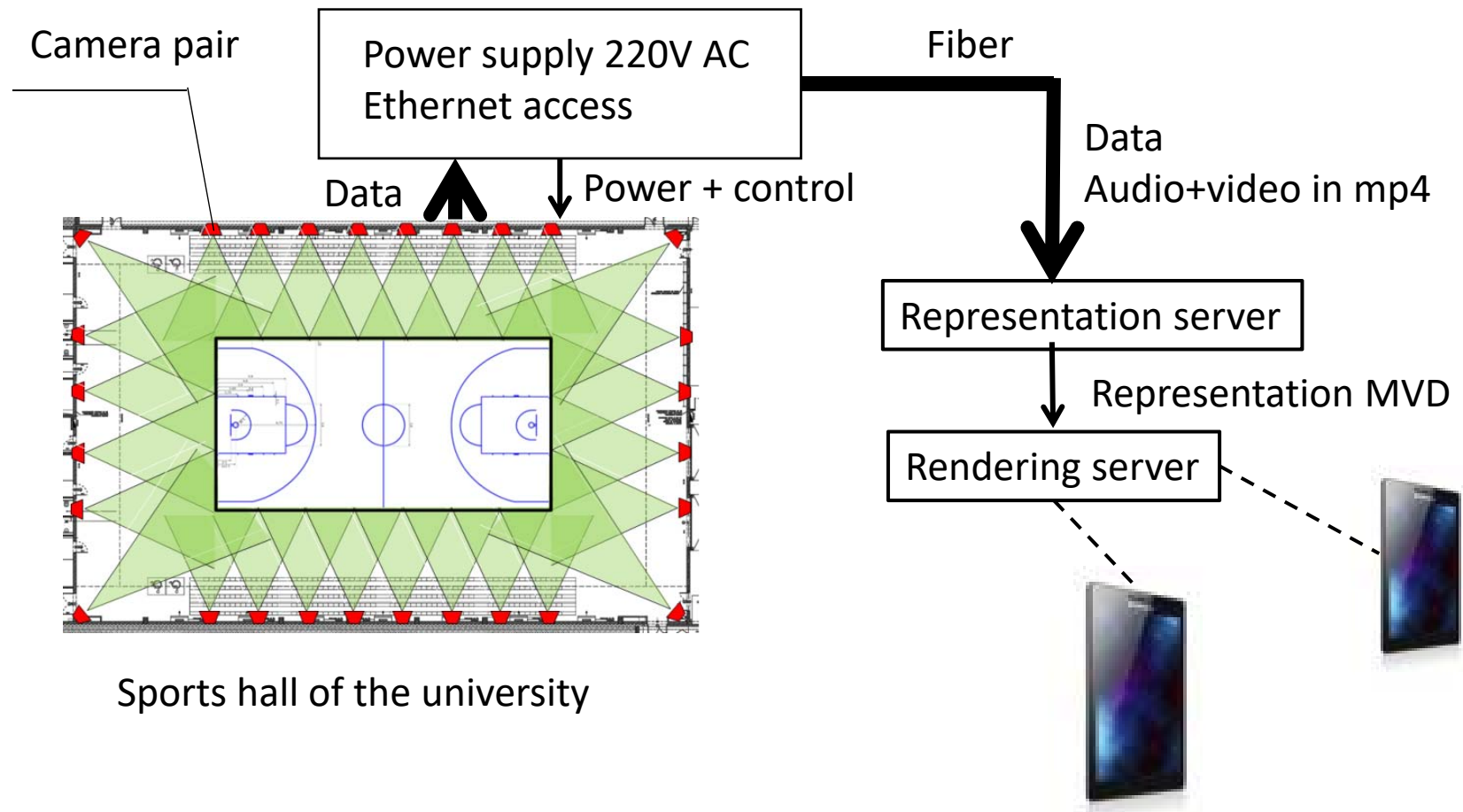
Experimental video acquisition system



- Camera pairs on fixed rig (~ 4.5 meter over floor)
- GoPro cameras



FTV system



System calibration



Wide-angle lenses : Reduced number of cameras but strong geometrical distortions



Virtual navigation example

Some processing errors !!



Virtual navigation example

Some processing errors !!



Virtual navigation example

Some processing errors !!



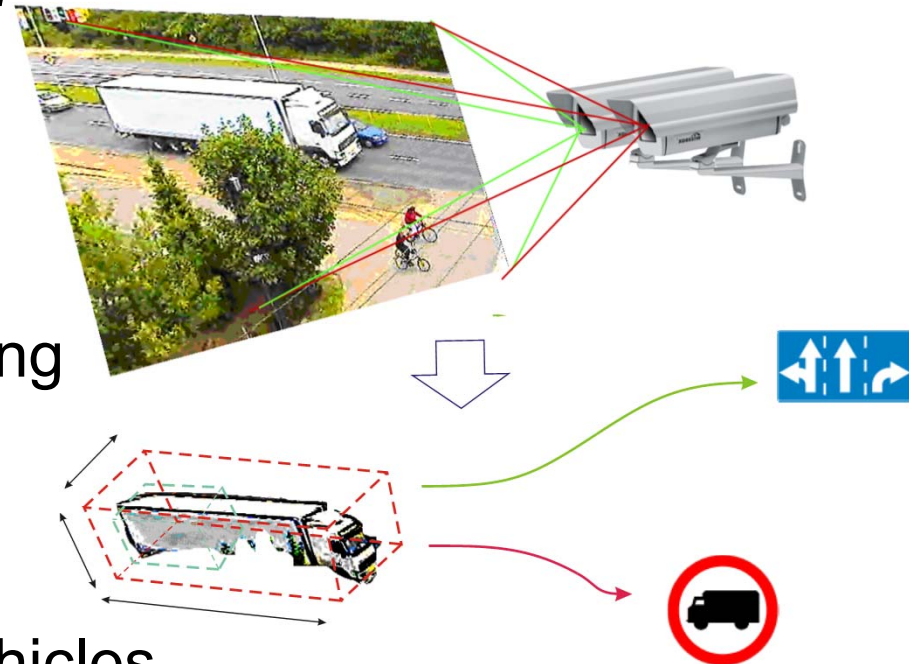
Pomiar rozmiarów pojazdów

„The automatic system for vehicle data extraction using stereoscopic video analysis”

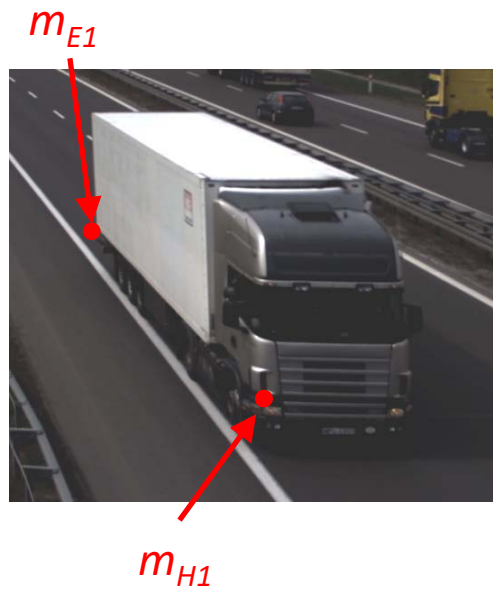
*Grant No NR02-0022-10/2011 founded by
The National Centre for Research and
Development (Poland)*

GOAL:

Design and develop system using stereoscopic video analysis in order to obtain information about vehicles, especially for estimating their size and classification of types of vehicles.



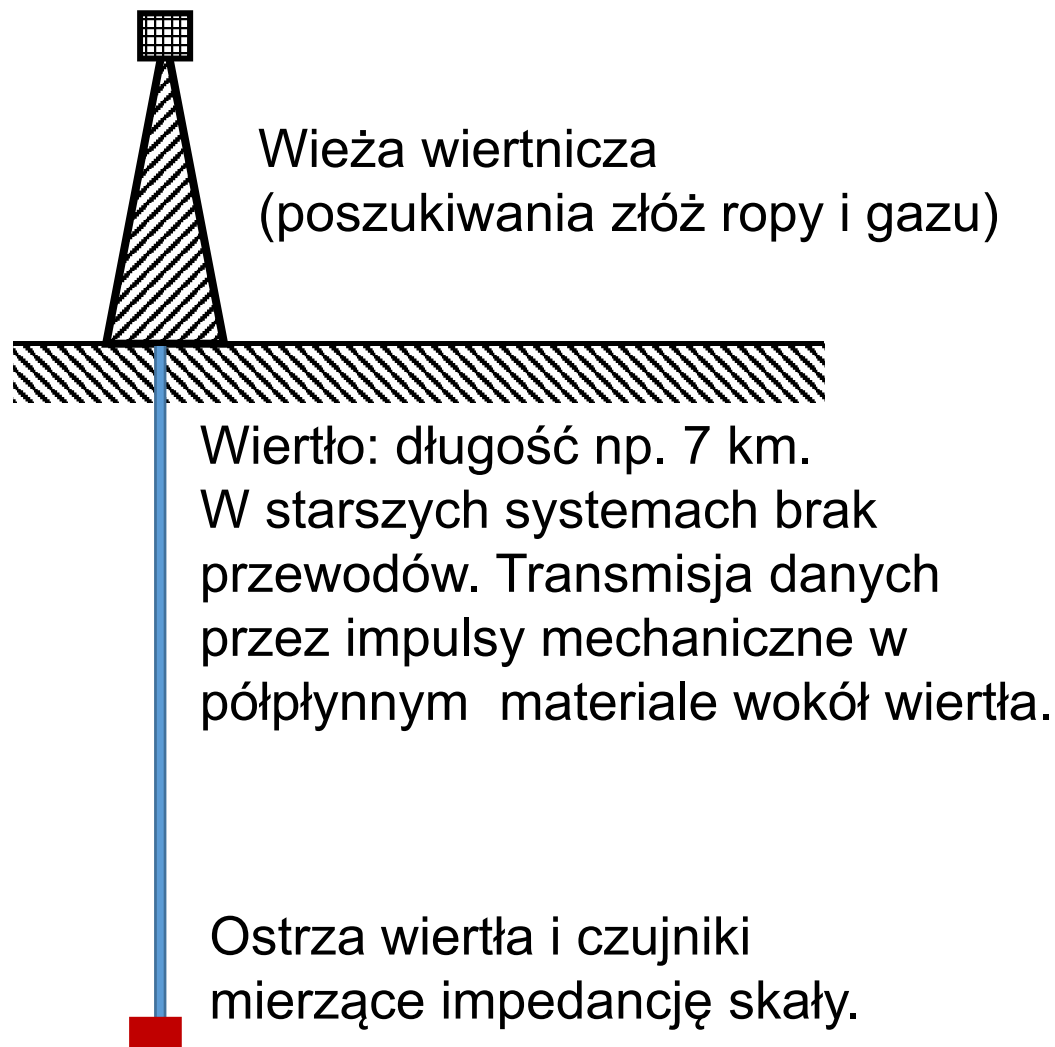
Pomiar rozmiarów pojazdów



Kompresja i transmisja danych w otworach wiertniczych

Zadanie:

Ponieważ koszt jednego odwiertu liczy się w milionach, steruje się kierunkiem wiercenia. Dane do podejmowania decyzji podejmuje się na podstawie badania impedancji skały.

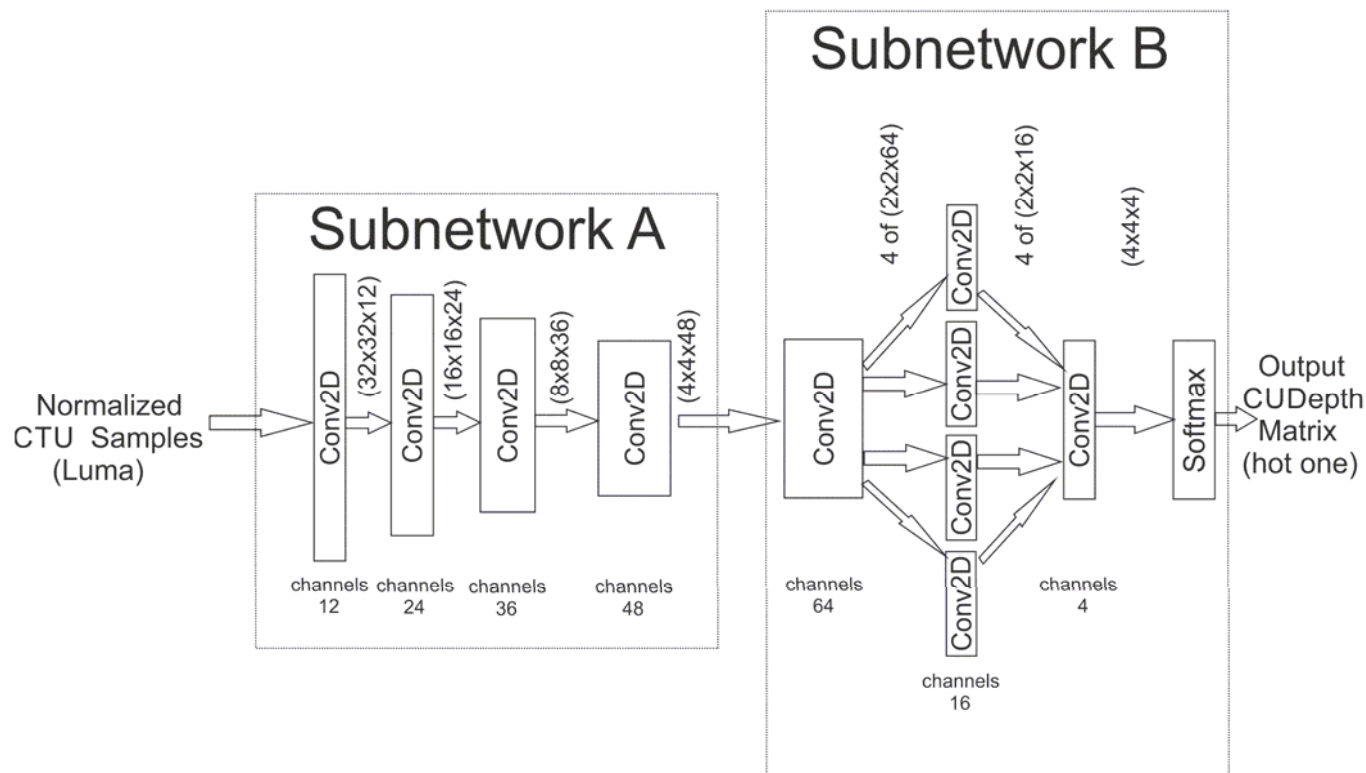


Inteligentne systemy dozoru wizyjnego

- Systemy automatycznie rozróżniające intruza :
Np. unikanie fałszywych alarmów od zwierząt.
- Systemy z klasyfikacją zachowań :
Wykrywanie bójek, omdleń, upadków itp.

Sterowanie koderem HEVC za pomocą sztucznych sieci neuronowych

- Wytrenowana na dużym zbiorze ramek z różnych filmów sztuczna sieć neuronowa steruje decyzjami dotyczącymi pracy kodera.

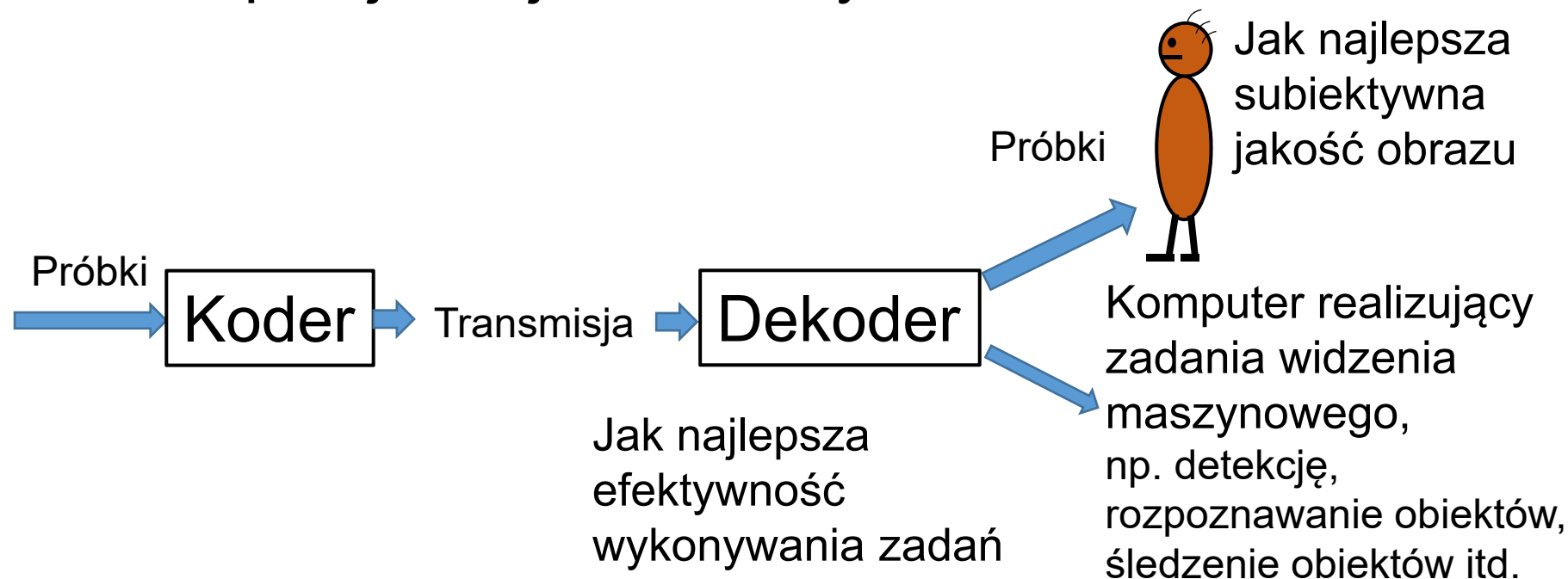


Kompresja wizji dla maszyn

- Tradycyjna kompresja:



- Kompresja wizji dla maszyn:



ICIP 2024 Panel on Video Coding for Machines

Abu Dhabi, October 29th, 2024



Shan Liu

Moderator: Shan Liu 2:30PM – 3:45PM, 2024.10.29



Marek
Domański



Jin Young
Lee



Dong Tian

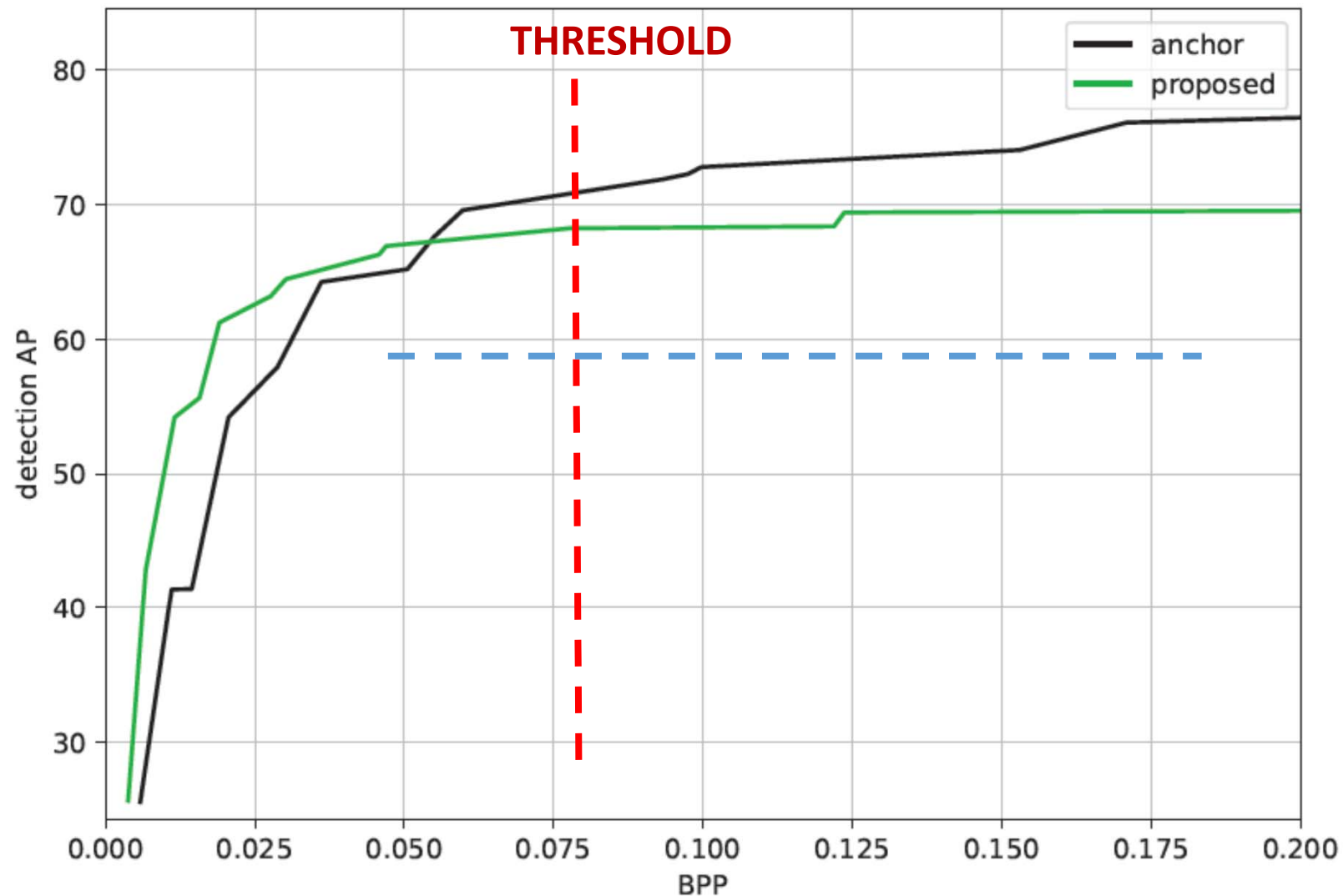


Giuseppe
Valenzise

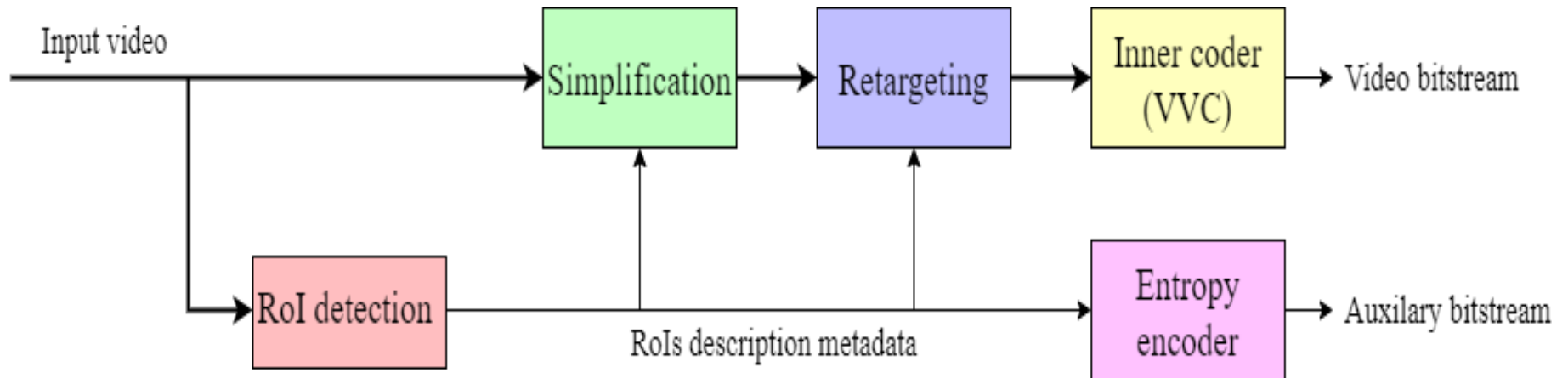


Honglei
Zhang

Rate – precision performance



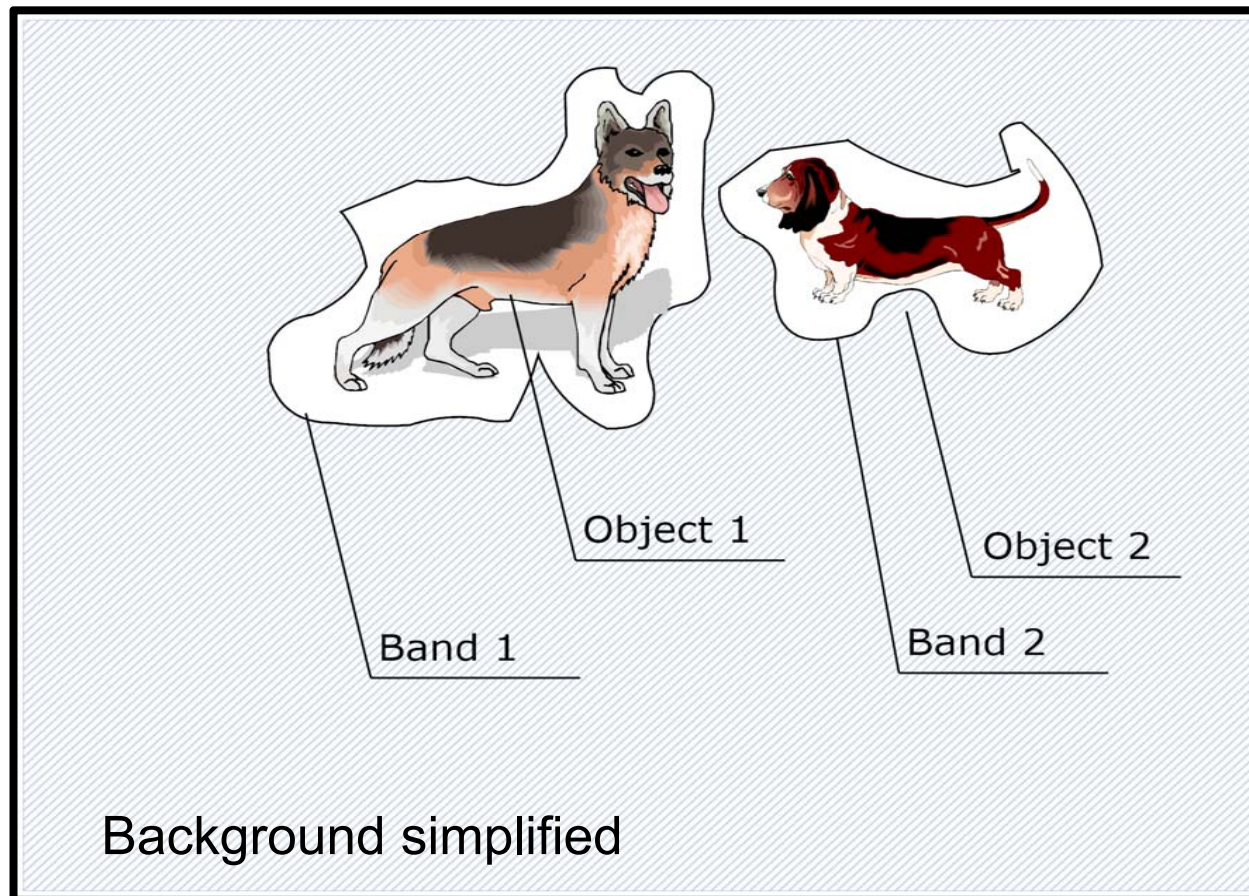
Propozycja zespołu PP



Preprocessing - Poznań proposal

Background simplification: - Lowpass filtering,
- Sample values downscaling

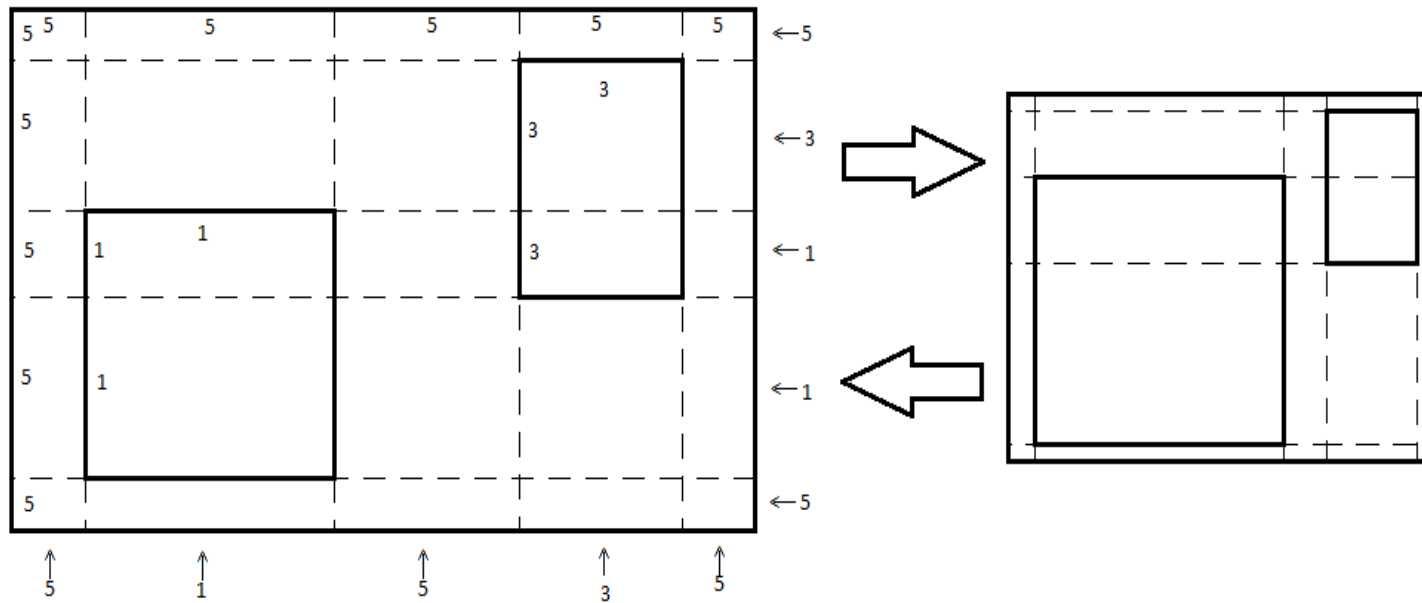
Image/ Video frame



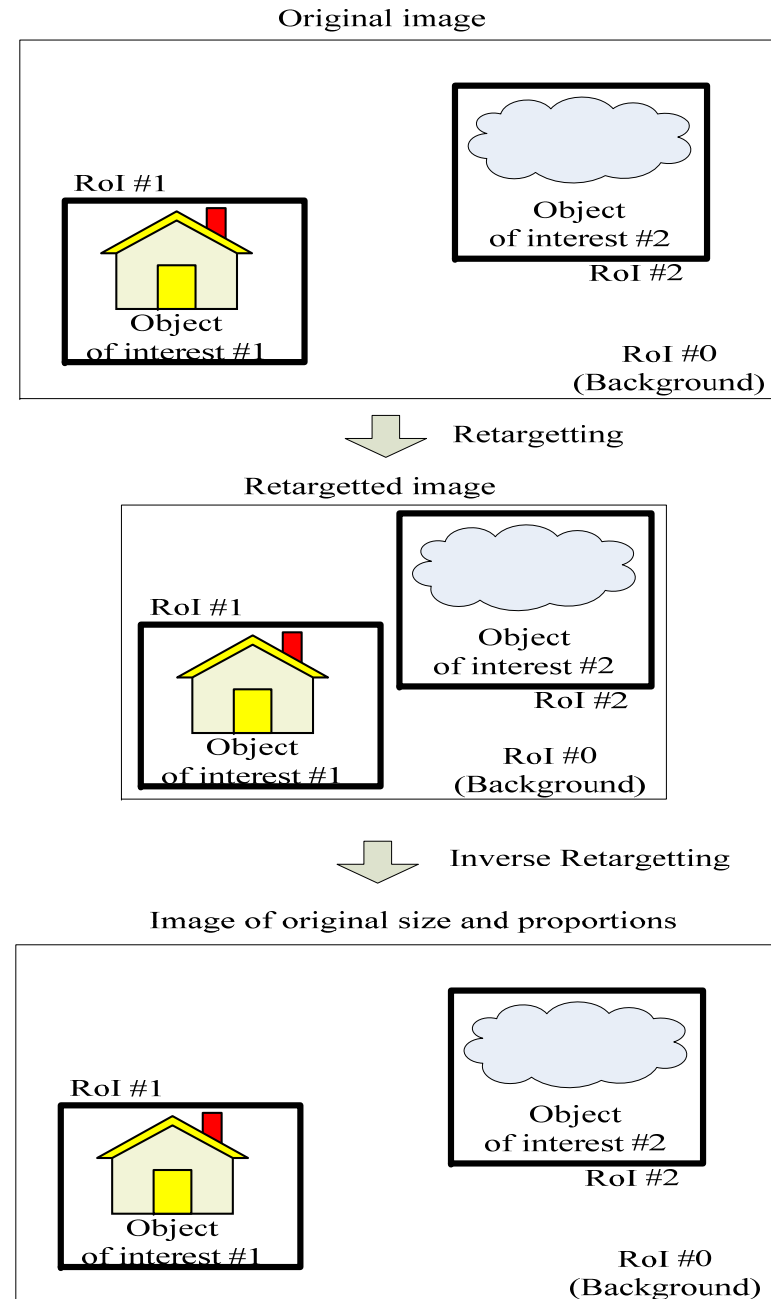
Simplification



Retargeting



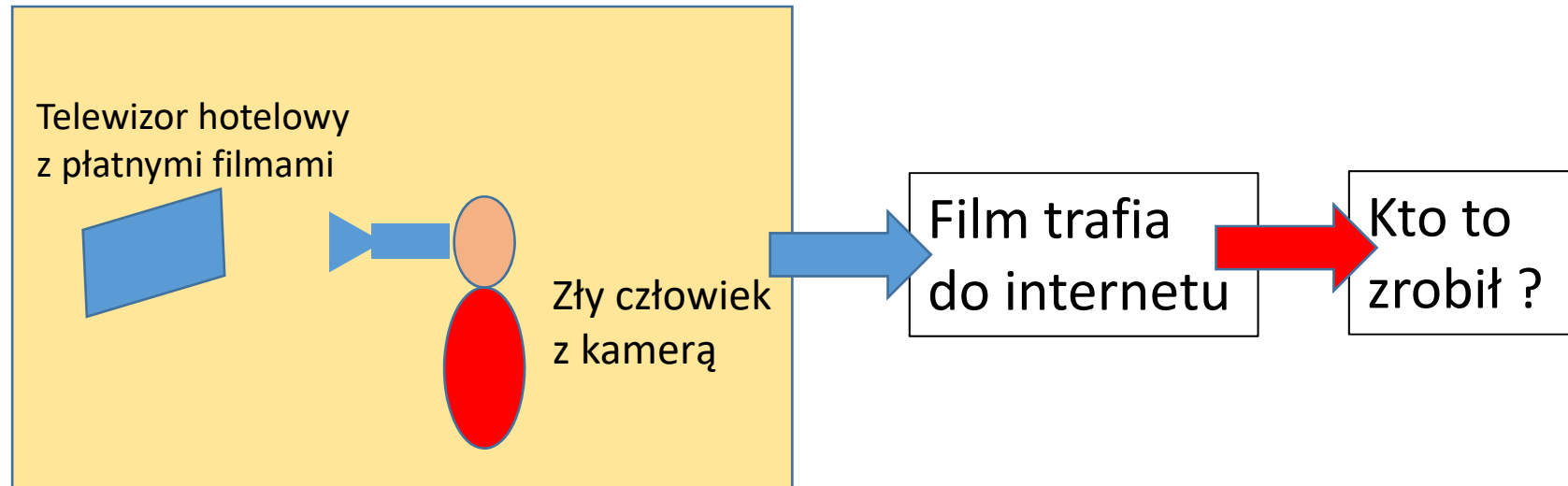
Poznań proposal – inverse retargeting



Znakowanie wodne w telewizji hotelowej

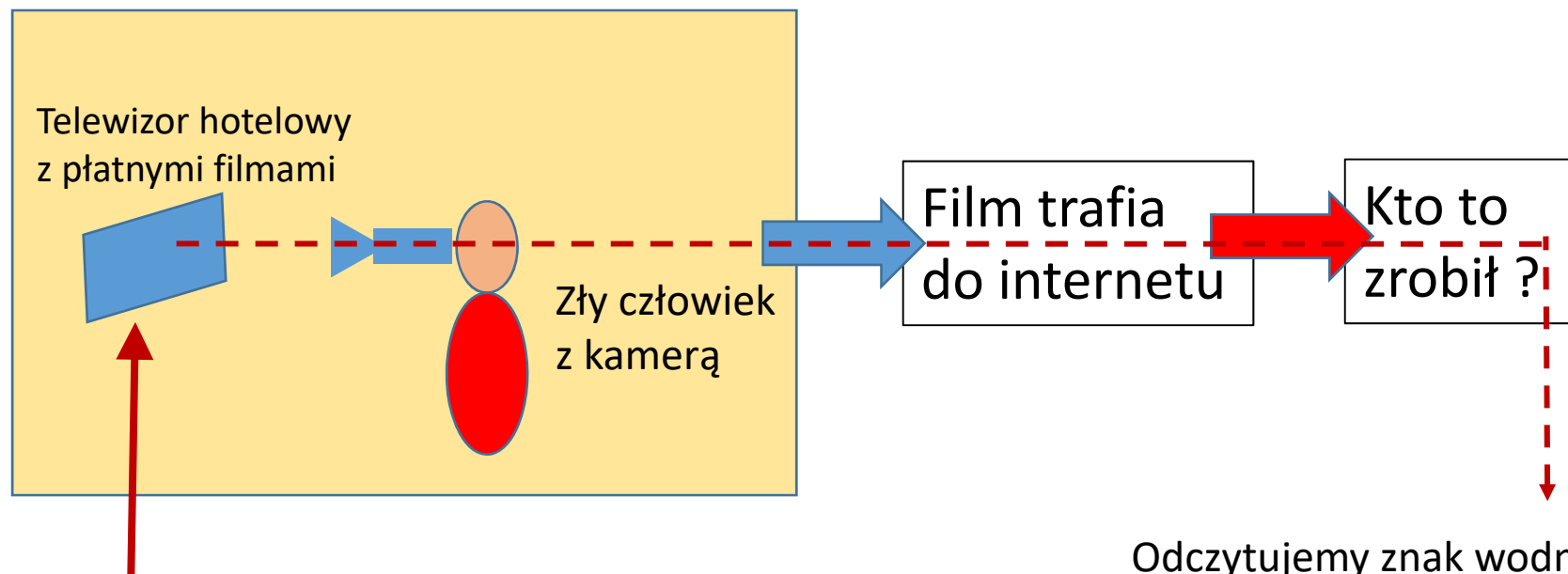
Problem w telewizji hotelowej

Pokój hotelowy



Znakowanie wodne

Pokój hotelowy



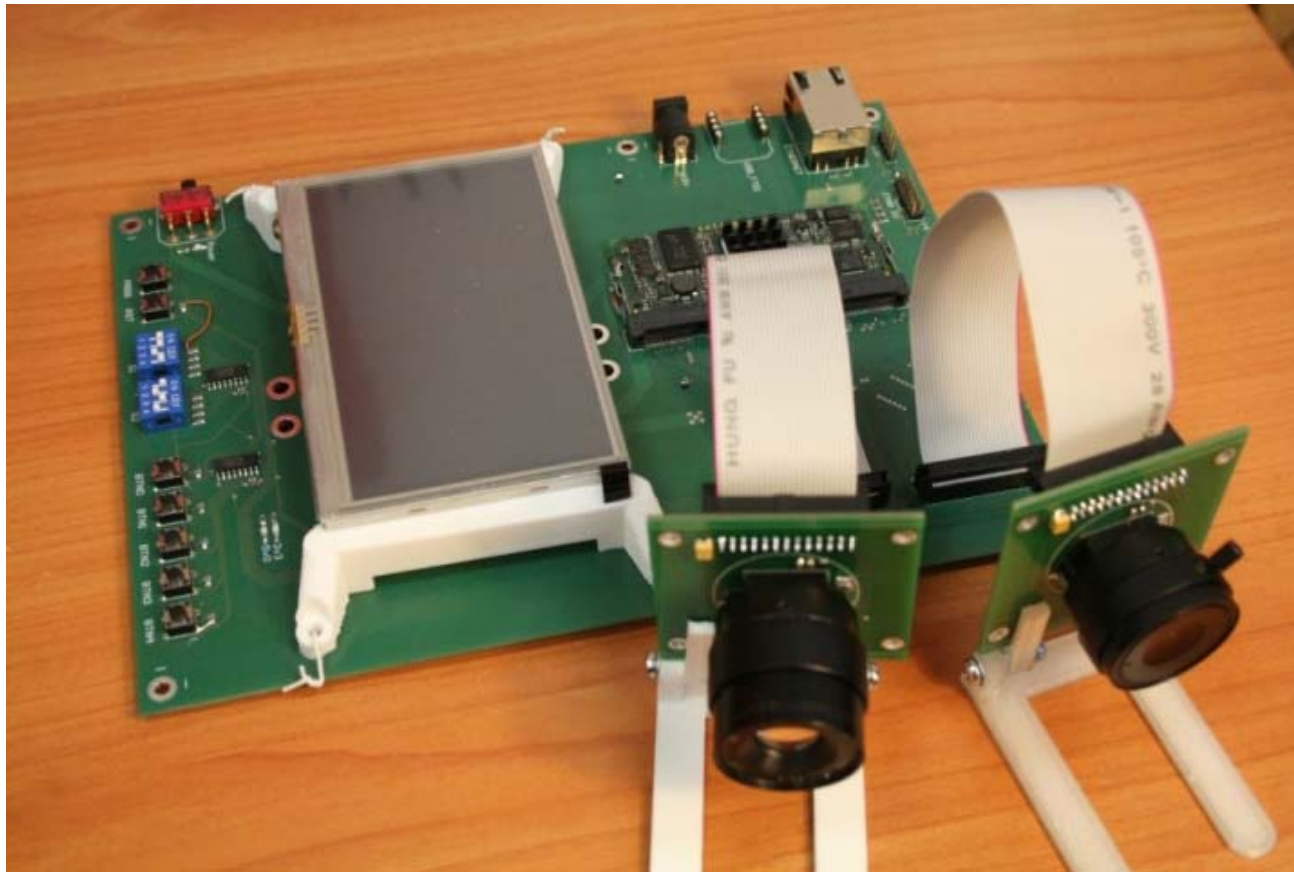
Dokładamy znak wodny:

- niewidoczny,
- trudny do usunięcia.

Odczytujemy znak wodny:

- data,
- nazwa hotelu,
- numer pokoju.

- **Katedra Telekomunikacji Multimedialnej i Mikroelektroniki**
 - **Wdrożenia przemysłowe**, rozwiązania sprzedane lub oferowane firmom.



Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

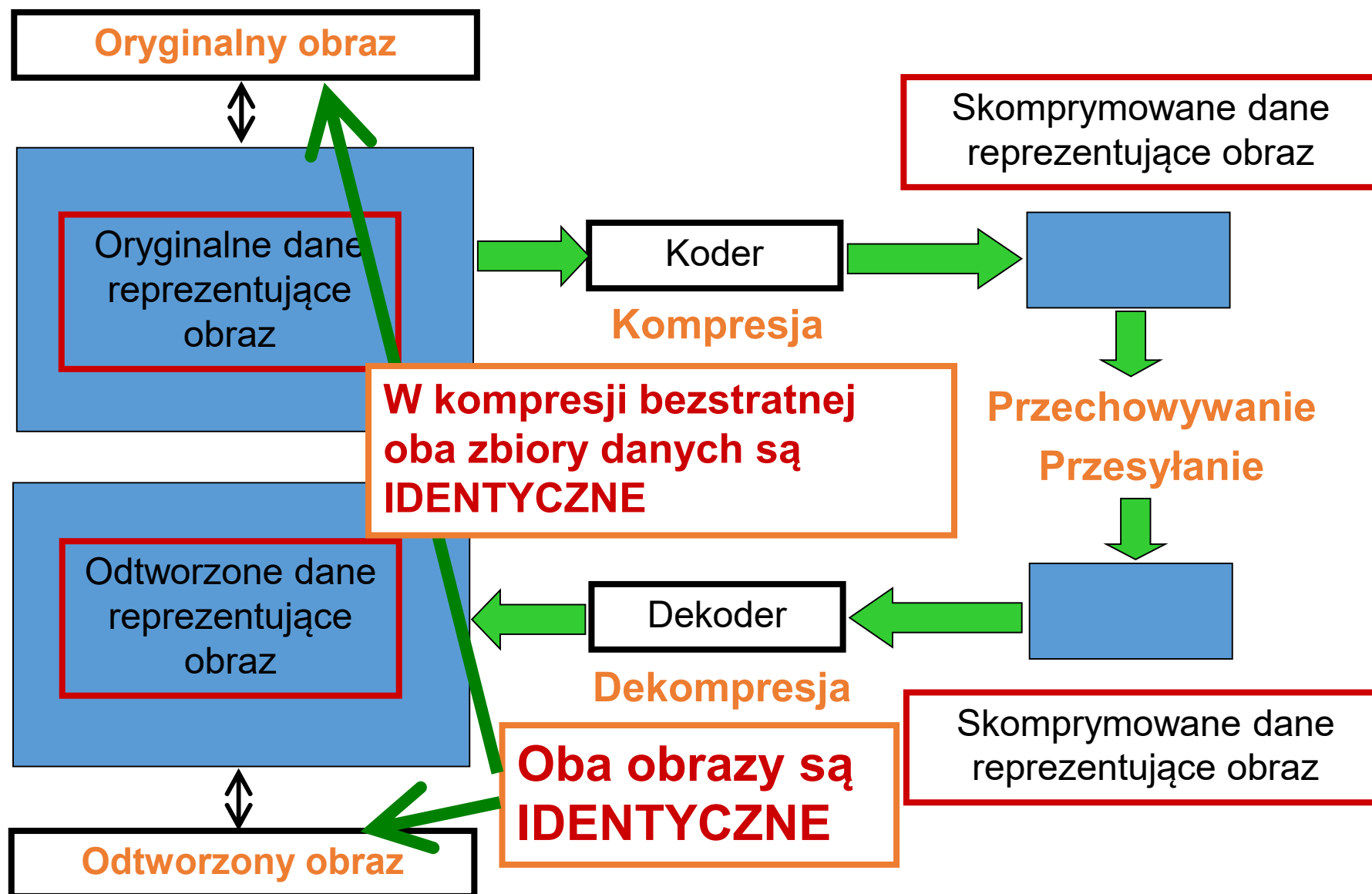
Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

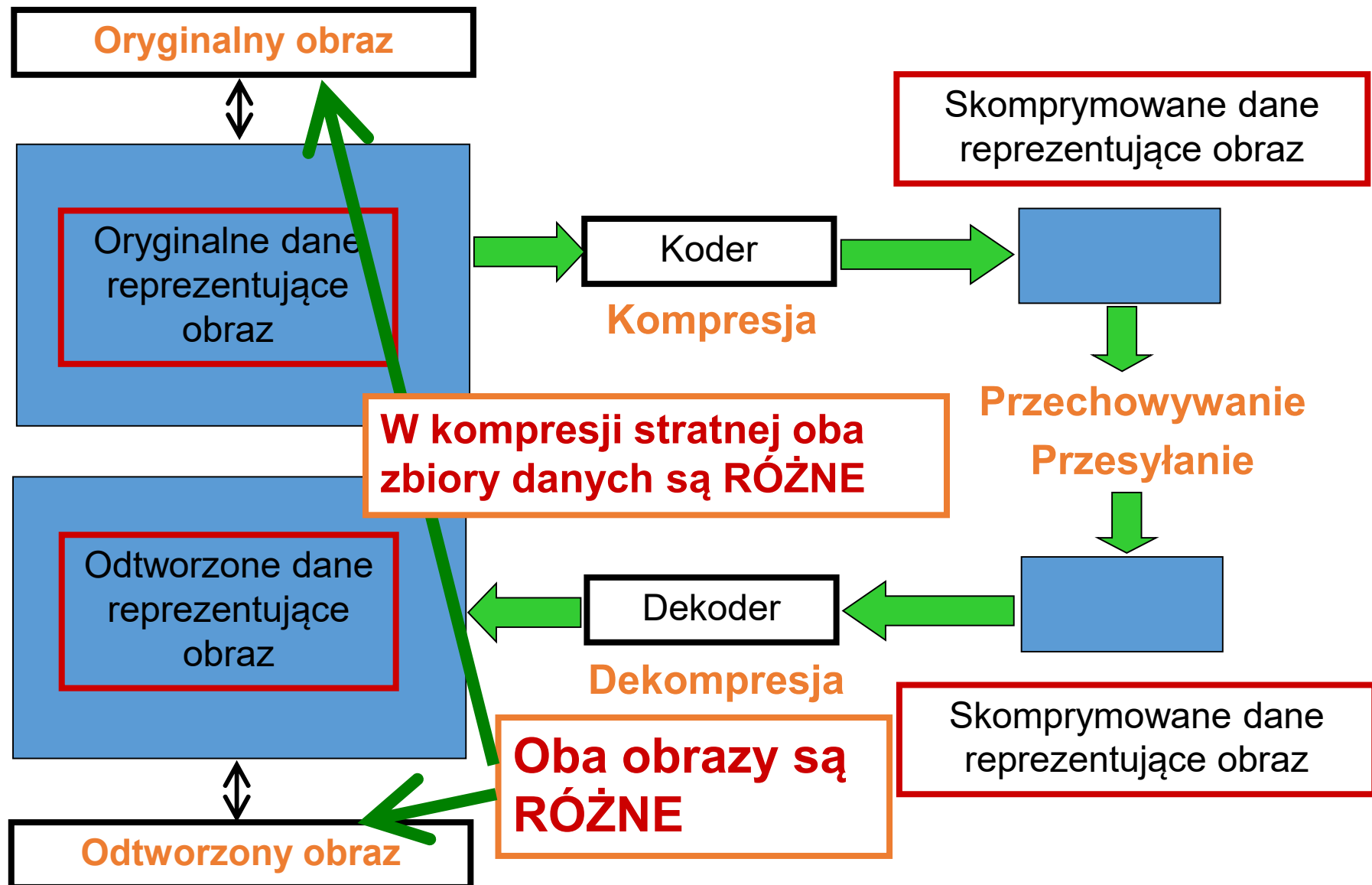
6. Kompresja

Wprowadzenie

Kompresja bezstratna (lossless compression) albo kompresja odwracalna (reversible compression)



Kompresja stratna (lossy compression) albo kompresja nieodwracalna (irreversible compression)



Miary kompresji obrazu

Współczynnik kompresji albo **stosunek kompresji** (*compression ratio*)

$$k_r = \frac{\text{długość oryginalnego łańcucha danych}}{\text{długość łańcucha danych po kompresji}} .$$

Ponadto często podaje się:

b_p – średnia liczba bitów na punkt obrazu [b/p]

$$b_p = \frac{\text{liczba bitów w reprezentacji obrazu}}{\text{liczba punktów w obrazie}}$$

Miary kompresji wizji i fonii

Dla skompresowanej wizji lub fonii :

Prędkość transmisji : kb/s, Mb/s, Gb/s.

Redukcja danych reprezentujących obrazy

Oryginalne dane reprezentujące cyfrowe obrazy i sygnały wizyjne charakteryzuje **nadmiarowość = redundancja**.

Informacje nadmiarowe

Redundancja przestrzenna

Redundancja czasowa

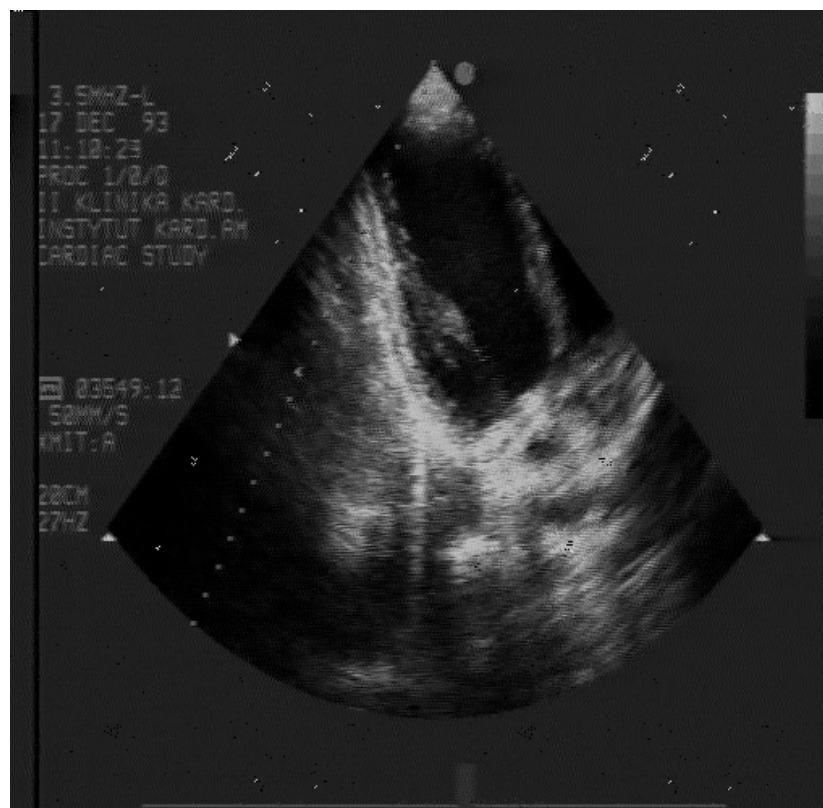
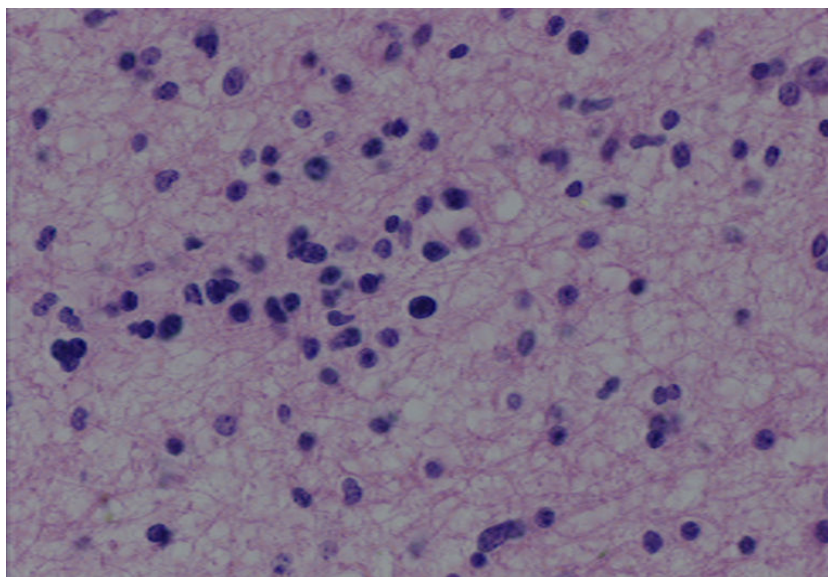
Redundancja kodowania

↔ (**redundancja statystyczna** - *statistical redundancy*)

Usuwanie informacji nadmiarowych pozwala na kompresję bezstratną.

Zastosowania kompresji bezstratnej

- Obrazy medyczne, zdjęcia geodezyjne itp.



Standardowe techniki bezstratnej kompresji obrazów

Wspólne cechy technik kompresji bezstratnej

1. Niezależna kompresja składowych (dowolnych) obrazów barwnych.
2. Kompresja składowych RGB jest w pełni odwracalna:
Typowe wartości współczynnika kompresji dla naturalnych obrazów monochromatycznych oraz obrazów barwnych o reprezentacji we współrzędnych RGB:
1,2 ÷ 2,5 – w zależności od treści obrazu i zastosowanej metody.
3. Kompresja połączona z transformacją barw (np. $RGB \rightarrow YC_R C_B$)
NIE jest odwracalna ze względu na błędy zaokrągleń transformacji barw.
Przy zastosowaniu transformacji barw $RGB \rightarrow YC_R C_B$ odpowiednie wartości współczynnika kompresji są większe o 20 ÷ 50% .
4. Nie ma możliwości regulacji wartości współczynnika kompresji.

Najważniejsze techniki kompresji bezstratnej

- **JPEG** ISO/IEC IS 10918, ITU-T Rec. T.81 (i inne)
Predykcja wewnątrzbrazowa nieadaptacyjna + Kodowanie Huffmana
(ew. arytmetyczne)
- **JPEG-LS** ISO/IEC IS 14495, ITU-T Rec. T.87, 870
Predykcja wewnątrzbrazowa adaptacyjna + Kodowanie wykładnicze
Golomba
- **GIF** (Graphics Interchange Format)
Kodowanie słownikowe obrazu
- **PNG** (Portable Network Graphics) ISO/IEC 15948
Predykcja wewnątrzbrazowa + Kodowanie słownikowe
+ Kodowanie Huffmana
- **TIFF** (Tagged Image File Format)
Format plików – przeważnie bez kompresji lub kompresja
bezstratna (algorytmy do wyboru: LZW (kodowanie słownikowe) i inne),
możliwa kompresja stratna (JPEG – w tej wersji format rzadko używany)

Porównanie

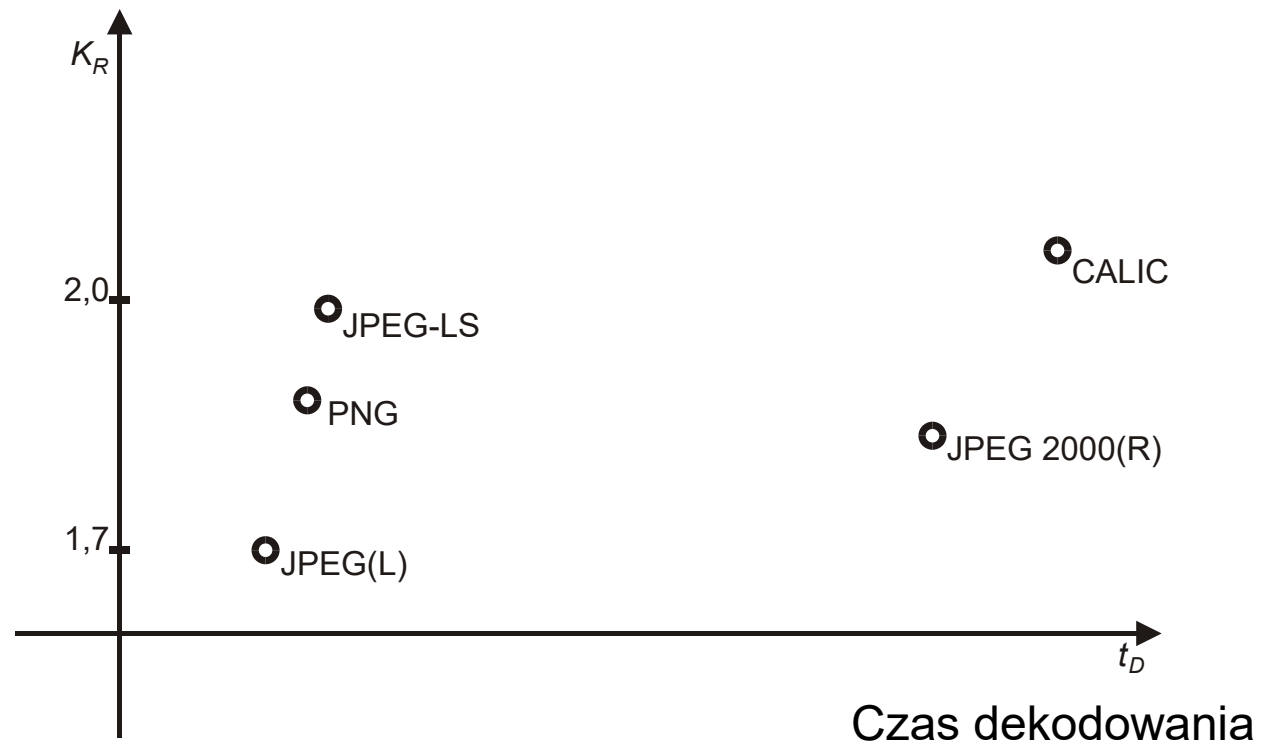
- Dla przykładowego testowego zbioru (obrazy naturalne, wykresy, złożone dokumenty):

JPEG-LS	3,19 bit/p
JPEG lossless Huffman	4,08 bit/p
JPEG lossless arithmetic	3,40 bit/p
PNG	3,46 bit/p

- JPEG-LS: kodowanie/dekodowanie szybsze niż dla PNG
- JPEG-LS: kodowanie/dekodowanie wolniejsze niż dla JPEG (tryb bezstratny) z kodowaniem Huffmana
- PNG wolny od patentów

Porównanie standardów kompresji bezstratnej

Średni współczynnik kompresji



Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

7. Stratna kompresja obrazów nieruchomych

7.1. Wprowadzenie

Redukcja danych reprezentujących obrazy

Przed wszystkim usuwa się

informacje nieistotne

Takie, których wykorzystanie nie jest uznane za konieczne w danym zastosowaniu.

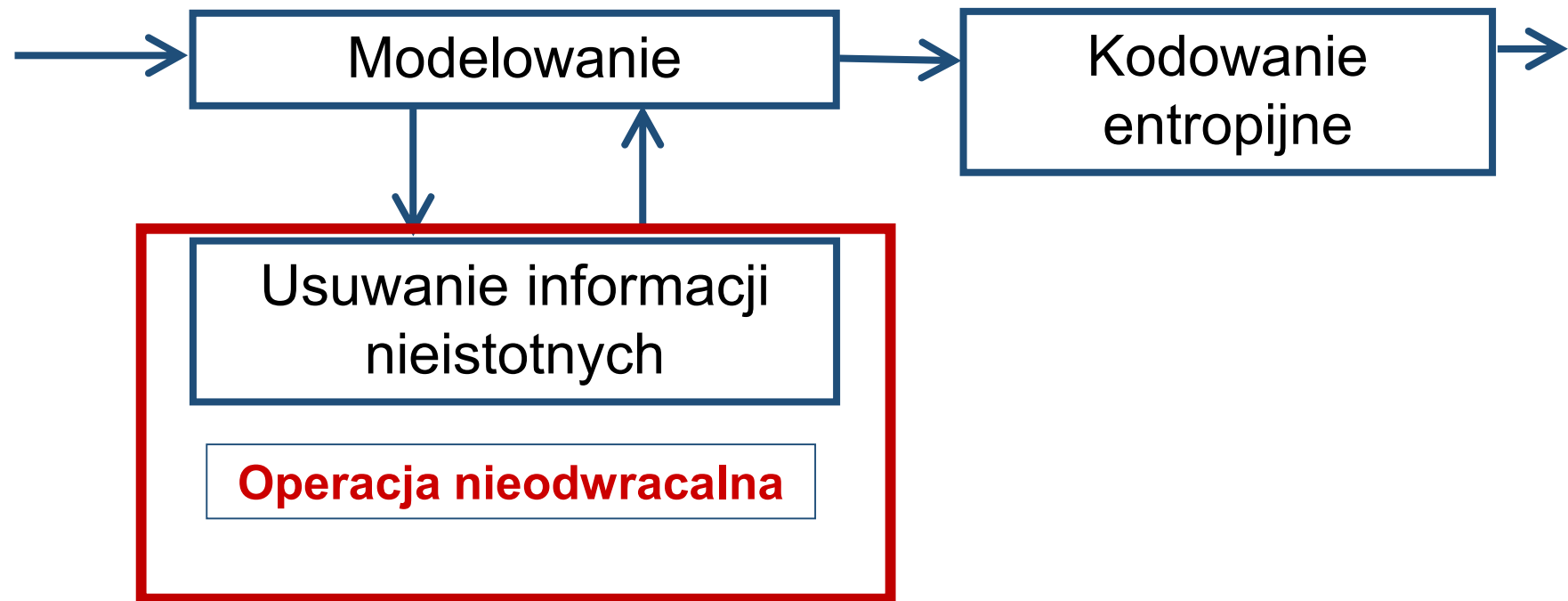
Wykorzystuje się właściwości ludzkiego systemu widzenia (np. usuwa się bardzo drobne szczegóły o małym kontraście).

Czyli usuwana jest

redundancja psychowizualna.

Dodatkowo usuwane są **informacje nadmiarowe**, jak w kompresji bezstratnej.

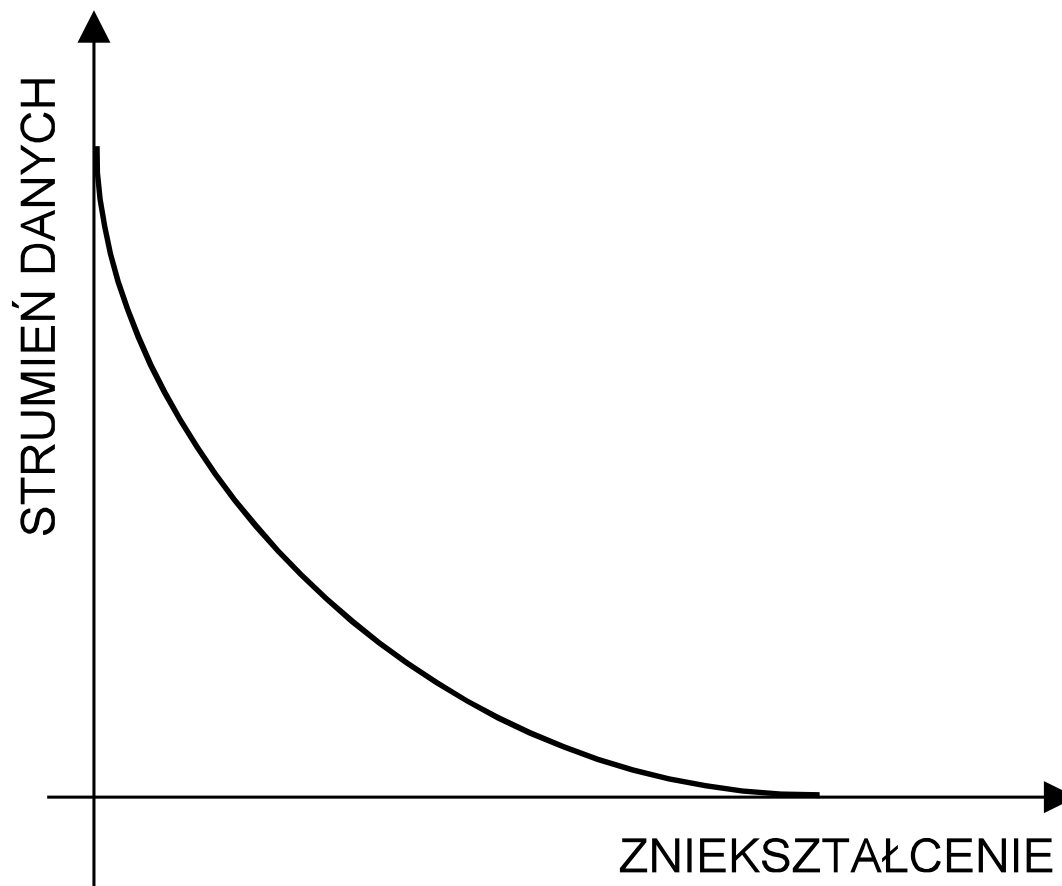
Kompresja stratna



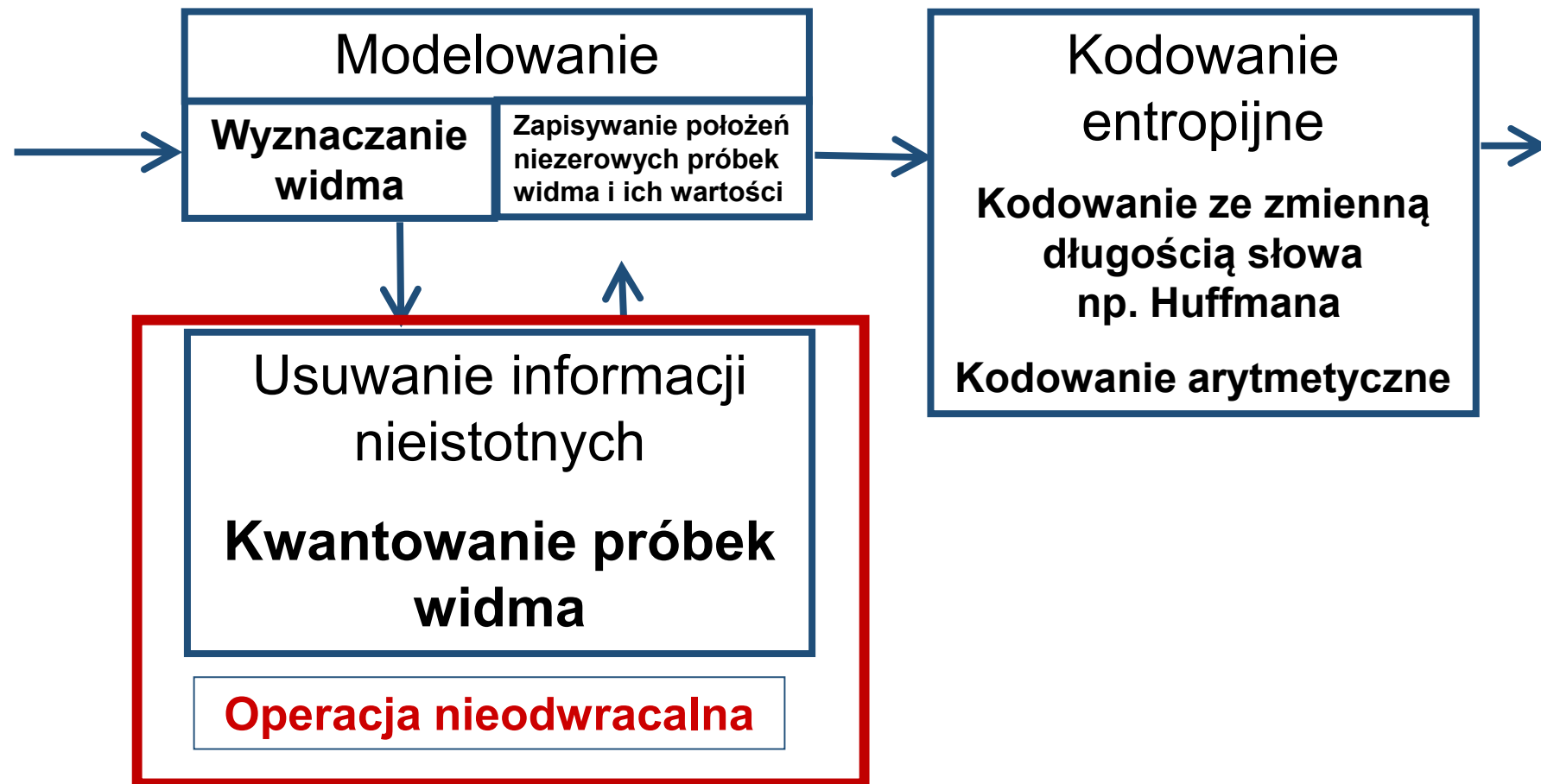
Usuwanie informacji nieistotnych (c.d.)

Klasyfikowanie coraz większej części informacji jako nieistotnej prowadzi do zwiększania kompresji, ale i do pogarszania jakości obrazów zdekodowanych.

Krzywa „strumień danych – zniekształcenie”



7.2. Idea kodowania transformatowego



Dyskretne przekształcenie Fouriera

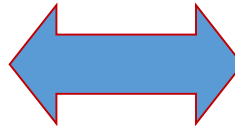
Przypomnienie:

DFT – *Discrete Fourier Transform*

$$F(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \cdot \exp(-j \frac{2\pi n k}{N}) \quad 0 \leq k, n \leq N-1$$

Dziedzina przestrzenna:

N próbek sygnału
okresowego



Dziedzina częstotliwości:

N próbek widma
okresowego

Dyskretne przekształcenie Fouriera

Discrete Fourier Transform DFT

okres sygnału obejmuje **N** próbek

próbki widma **zespolone**

Dyskretne dwuwymiarowe przekształcenie Fouriera - definicje

DFT – Discrete Fourier Transform

$$F(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \cdot \exp(-j \frac{2\pi n k}{N}) \quad 0 \leq k, n \leq N-1$$

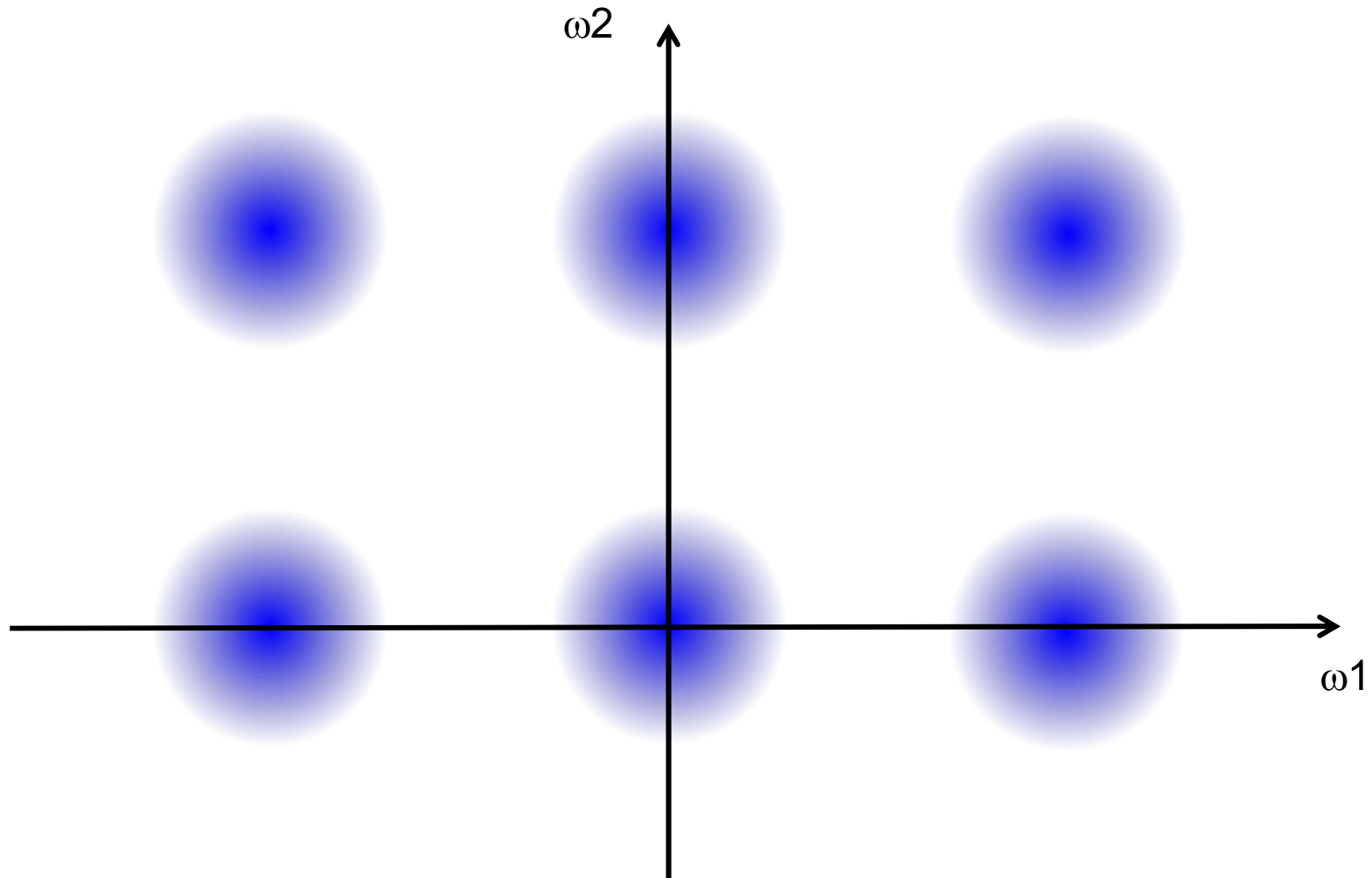
2-D DFT

$$F(k_1, k_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2) \cdot \exp\left(\frac{-j2\pi n_1 k_1}{N_1} + \frac{-j2\pi n_2 k_2}{N_2}\right)$$

$$0 \leq k_1, n_1 \leq N_1-1, \quad 0 \leq k_2, n_2 \leq N_2-1$$

$$f(n_1, n_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F(k_1, k_2) \cdot \exp\left(\frac{j2\pi n_1 k_1}{N_1} + \frac{j2\pi n_2 k_2}{N_2}\right)$$

Widmo okresowe



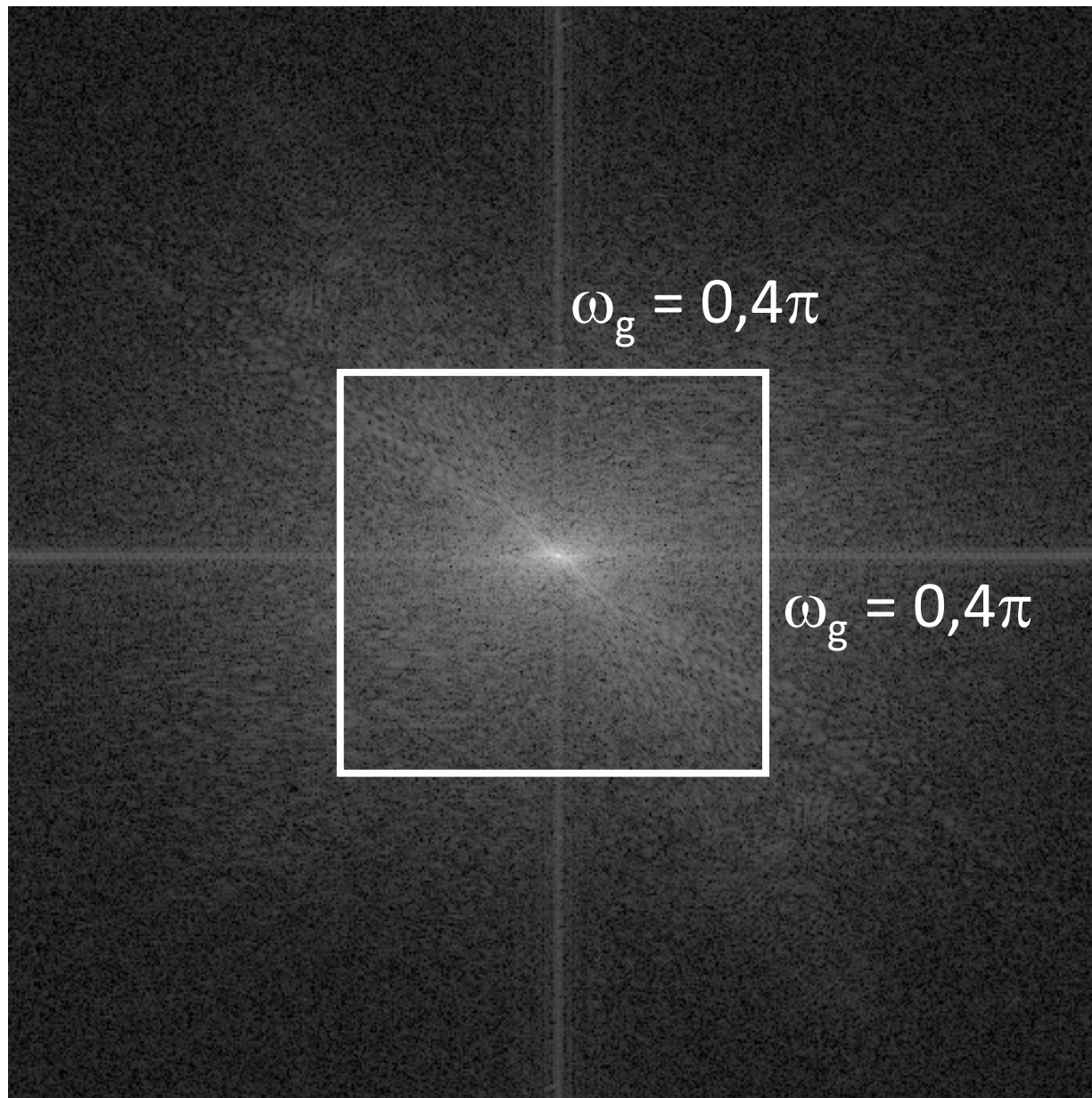
Prymitywne kwantowanie próbek widma

- Usuwamy próbki widma odpowiadające pulsacjom $|\omega_1| > \omega_g$, $|\omega_2| > \omega_g$.



Oryginał

Widmo amplitudowe



Filtracja
dolnoprzepustowa
 $\omega_g = 0,4\pi$

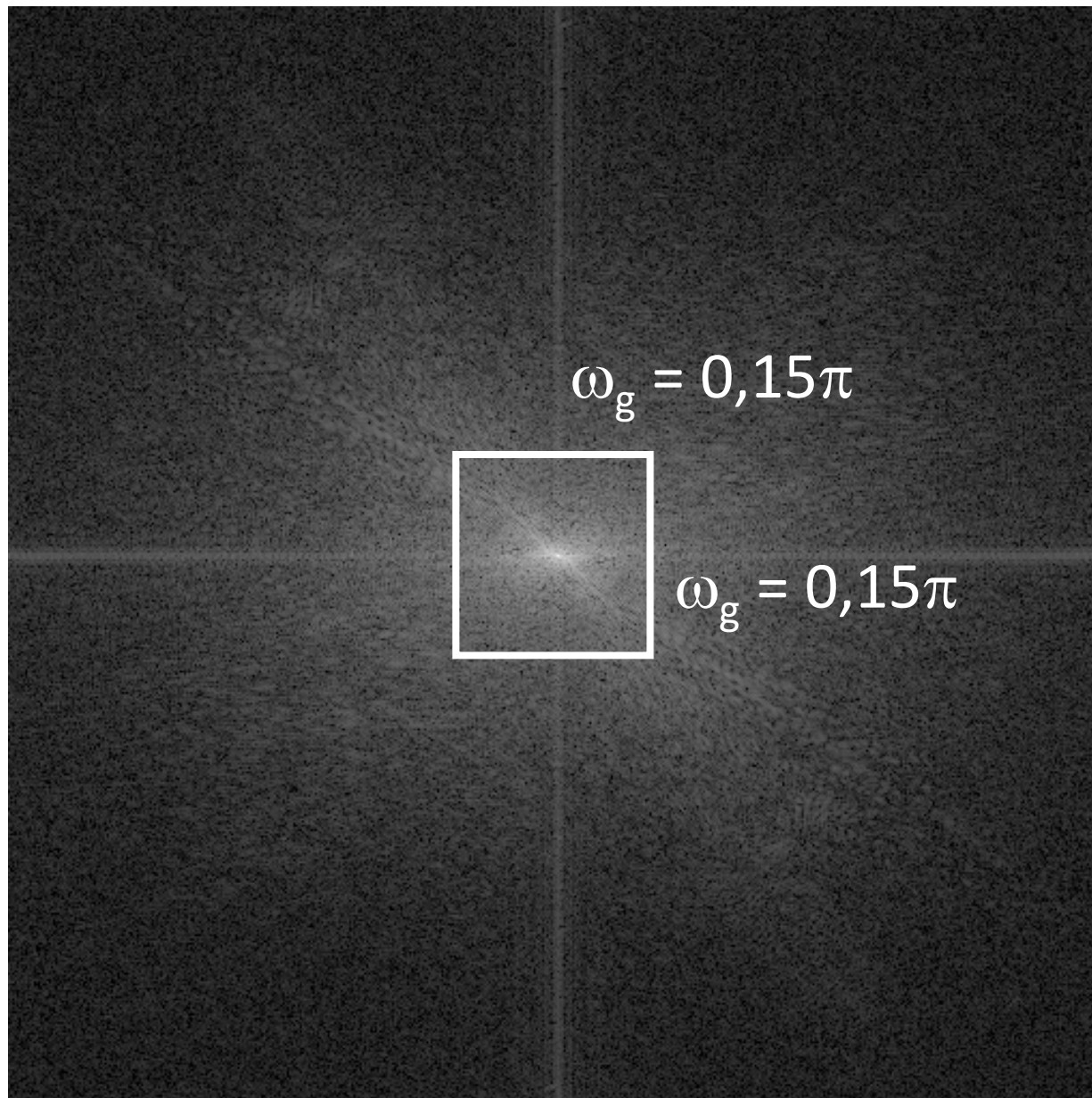
Pozostawiamy
16%
próbek DFT



$$\omega_g = 0,4\pi$$

Obraz
odtworzony
z 16%
próbek DFT

Widmo amplitudowe



Filtracja
dolnoprzepustowa
 $\omega_g = 0,15\pi$

Pozostawiamy
2,25%
próbek DFT



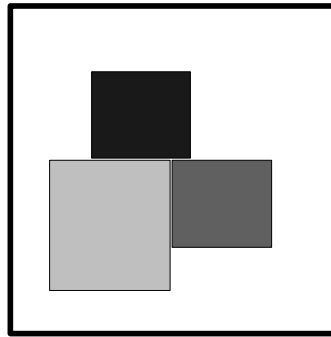
$$\omega_g = 0,15\pi$$

Obraz
odtworzony
z 2,25%
próbek DFT

7.3. Praktyczna realizacja kodowania transformatowego

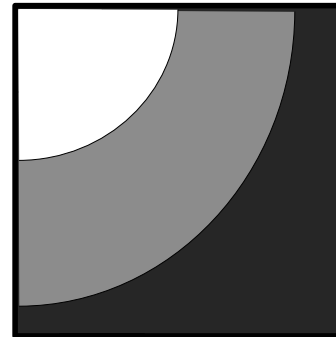
1. Obraz jest dzielony na bloki o rozmiarze 8×8 punktów.

2. Każdy blok jest poddawany transformacji



Blok obrazu

Obszary jasne i ciemne
rozłożone przypadkowo.



Blok transformaty

Obszary jasne i ciemne
prawie uporządkowane.

3. Próbkki transformaty są kwantowane (zaokrąglane)

Dużo próbek ma zerową wartość.

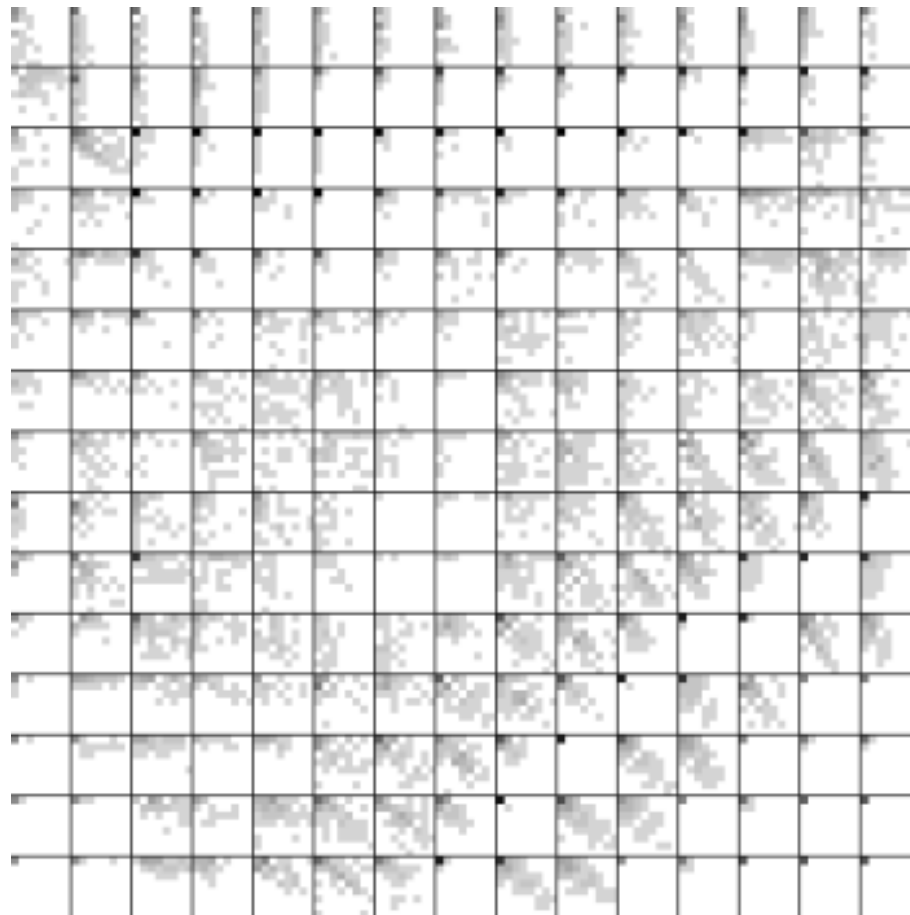
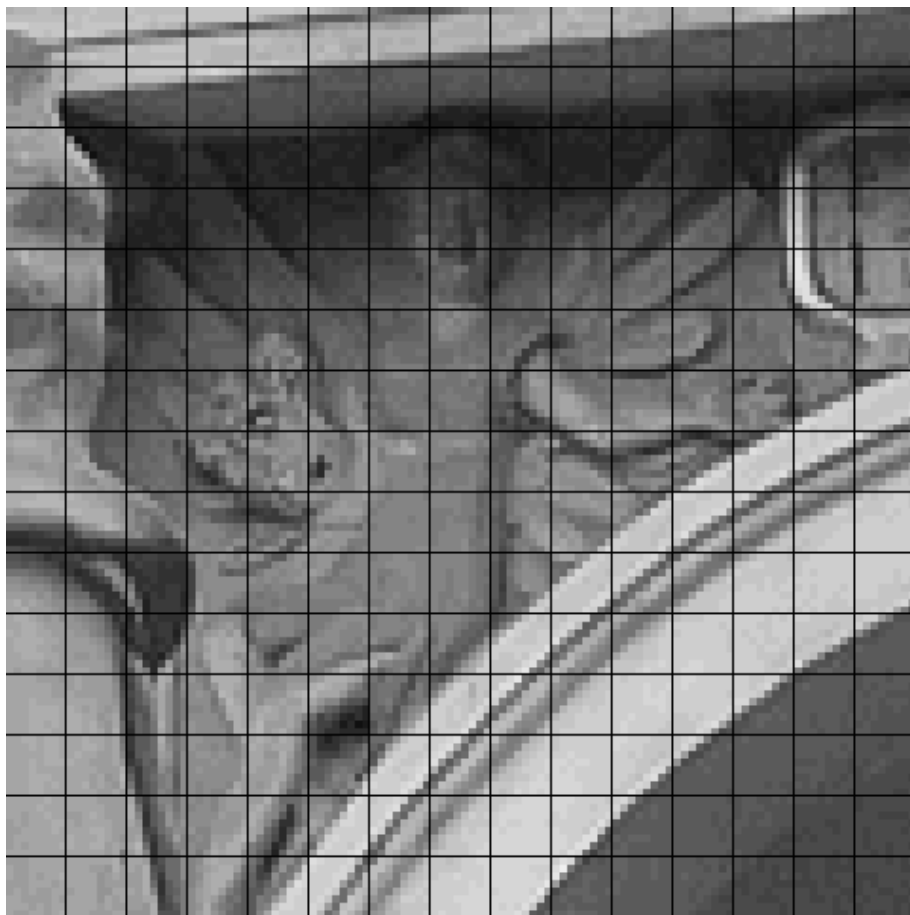
4. Koduje się liczby określające :

- odległości między niezerowymi próbkami,
- wartości niezerowych próbek.

Stosuje się kody o zmiennej długości słowa.

Widmo amplitudowe bloków 8 x 8 rzeczywistego obrazu

Ciemniejsze punkty
– większe wartości próbek widma



7.4. Dyskretne przekształcenie kosinusowe

Za pomocą DFT obliczmy próbki
widma sygnału okresowego



Na styku kolejnych
okresów obrazu
powstają skokowe
zmiany wartości.



Składowe odpowiadające
dużym częstotliwościom
mają znaczne amplitudy.



Usuwanie tych składników
widma wynikających
z konstrukcji sygnału
okresowego powoduje
spore straty jakości.

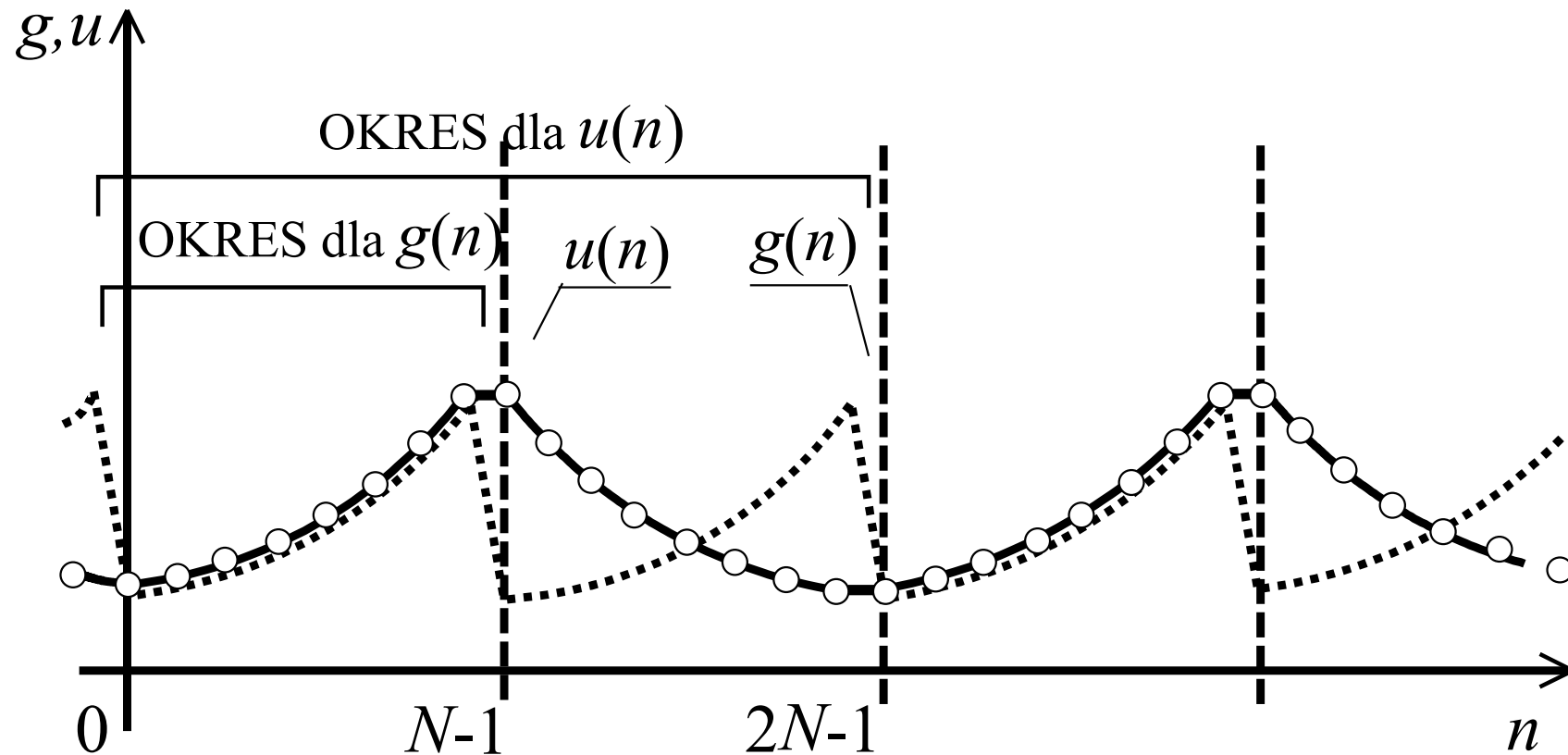
**Może inaczej konstruować sygnał okresowy
dwuwymiarowy ?**



Na przykład tak.



Alternatywny sposób konstrukcji sygnału okresowego objaśniony dla sygnału jednowymiarowego



- Nowy sygnał **okresowy** $u(n)$ (linia ciągła) ma okres składający się z $2N$ próbek.
- Dla tego nowego sygnału okresowego $u(n)$ wyznaczymy teraz dyskretną transformatę Fouriera.
- ***Podręcznik „Obraz cyfrowy” – str. 101 i następne***

Sygnał okresowy dwuwymiarowy



Dyskretne przekształcenie kosinusowe (*Discrete Cosine Transform*)

Przekształcenie proste

$$U(k) = \sqrt{\frac{2}{N}} \cdot C(k) \sum_{n=0}^{N-1} u(n) \cdot \cos \frac{\pi(2n+1)k}{2N}, \quad k = 0, \dots, N-1.$$

$$C(k) = \frac{1}{\sqrt{2}} \quad \text{dla } k = 0 \text{ oraz } C(k) = 1 \quad \text{dla } k \neq 0.$$

Przekształcenie odwrotne

$$u(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} C(k) \cdot U(k) \cdot \cos \frac{\pi(2n+1)k}{2N}, \quad n = 0, \dots, N-1.$$

Dyskretne przekształcenie kosinusowe (c.d.)

Dwuwymiarowa dyskretna transformata kosinusowa jest separowalna.

Przekształcenie 8-punktowe (wierszami lub kolumnami):

$$F(k) = \frac{1}{2} C(k) \sum_{n=0}^7 f(n) \cdot \cos[(2n+1)k\pi / 16] ,$$

gdzie $n = 0, \dots, 7$ jest numerem próbki obrazu $f(n)$,

$k = 0, \dots, 7$ numerem współczynnika transformaty,

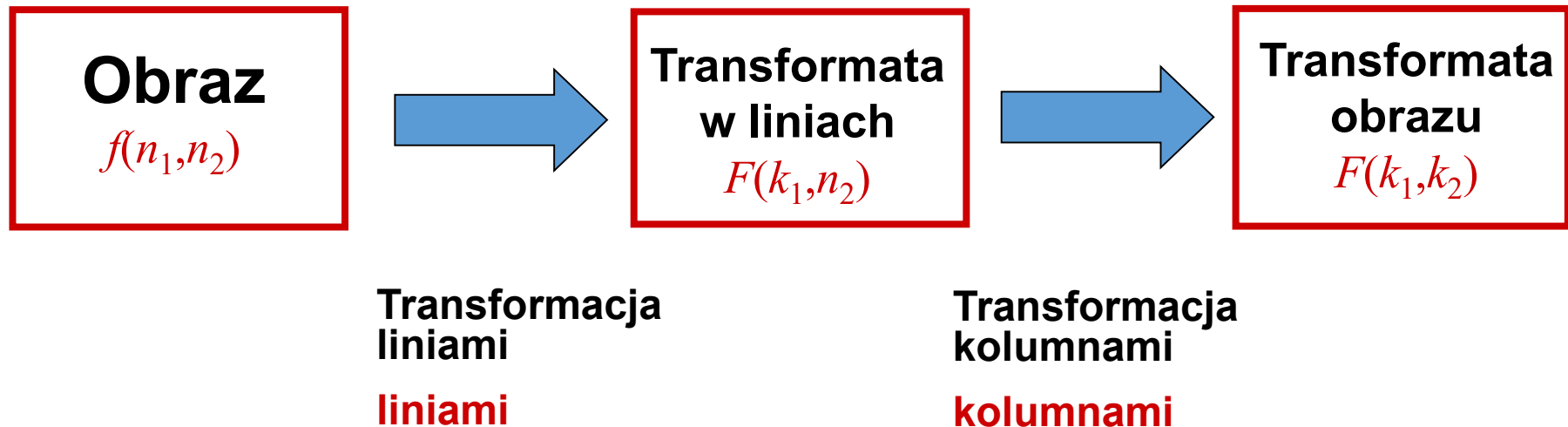
$C(k) = 1, k \neq 0$ i $C(0) = \sqrt{0,5}$.

Obliczanie DCT

$$U(k_1, k_2) = \frac{2}{\sqrt{N_1 N_2}} \cdot C(k_1) \cdot C(k_2) \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} u(n_1, n_2) \cdot \cos \frac{\pi(2n_1 + 1)k_1}{2N_1} \cdot \cos \frac{\pi(2n_2 + 1)k_2}{2N_2}$$

$$0 \leq k_1, n_1 \leq N_1-1, \quad 0 \leq k_2, n_2 \leq N_2-1$$

Separowalna !!!



Wynik: $F(k_1, k_2)$.

Dyskretne przekształcenie kosinusowe (c.d.)

Złożoność obliczeniowa

(8×8) – punktowa DCT liczona według definicji:

Wyznaczenie 4096 iloczynów próbek, dwóch wartości kosinusów i współczynników skalujących.

Następnie należy dodać odpowiednie składniki.

Szybkie algorytmy DCT:

Przekształcenie jednowymiarowe 8-punktowe:

11 ÷ 13 mnożeń i 29 dodawań,

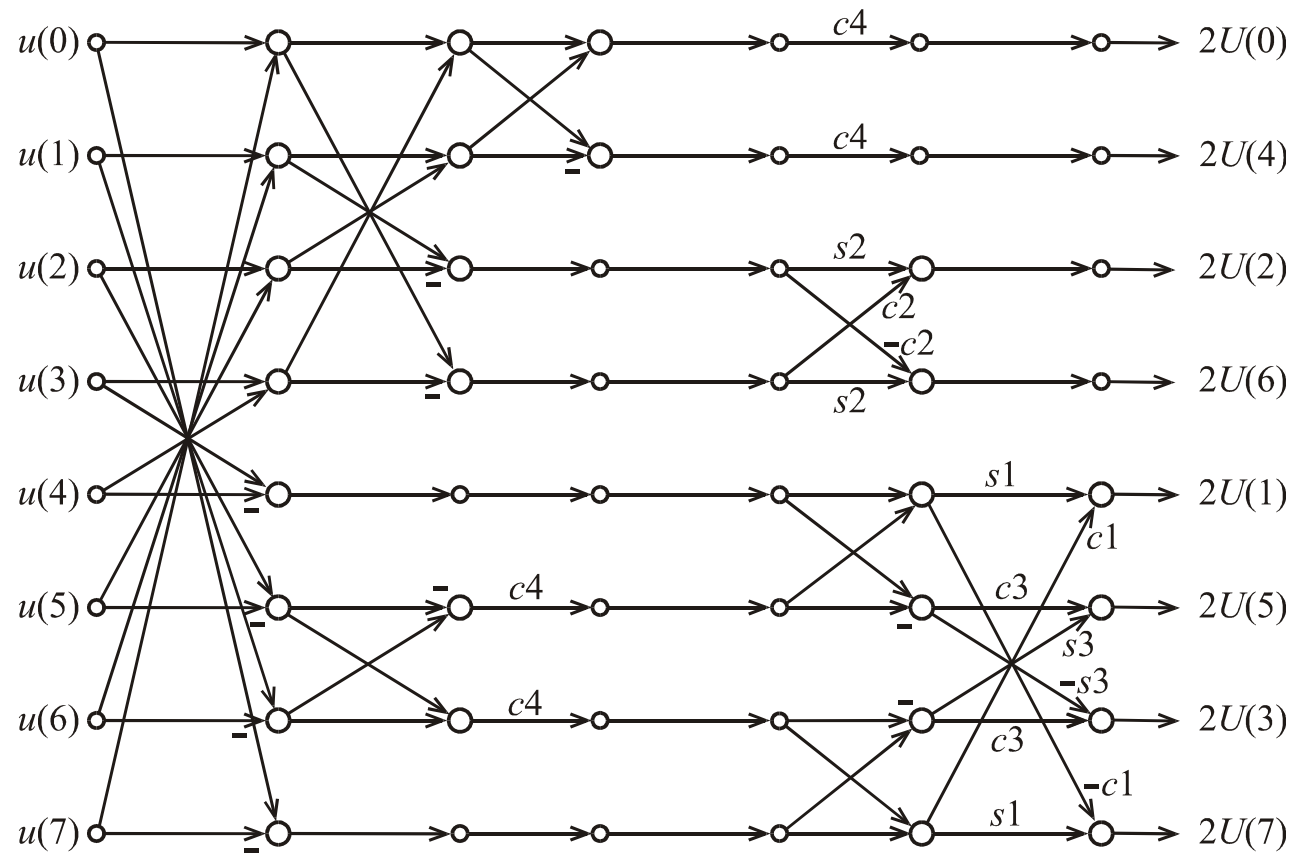
transformacja przeskalowana wymaga tylko 5 mnożeń i 29 dodawań.

Przekształcenie dwuwymiarowe (8×8) –punktowe:

Dalsze zmniejszenie liczby operacji, ale nieregularne algorytmy.

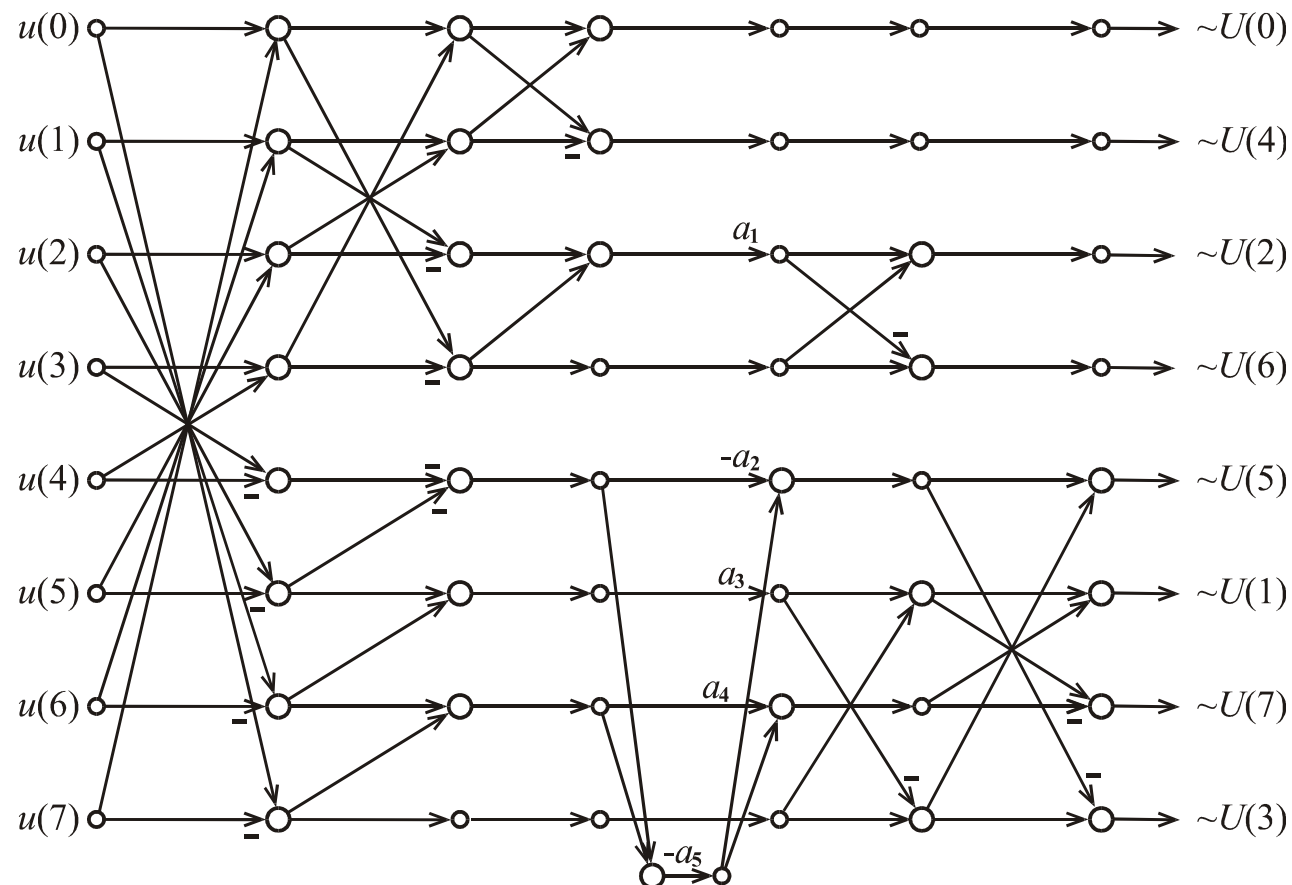
mnożeń ok. 100-120, dodawań 430-460.

Szybkie przekształcenie kosinusowe



$$c_i = \cos \frac{\pi i}{16}, s_i = \sin \frac{\pi i}{16}$$

Szybkie przekształcenie kosinusowe



$$a_1 = 0,707, a_2 = 0,541, a_3 = 0,707, a_4 = 1,307, a_5 = 0,383$$

Problem

- Błędy zaokrągleń w koderze i w dekodерze

7.5. Kwantowanie współczynników transformaty kosinusowej

Najpierw od wartości wszystkich próbek odejmowana jest liczba 2^{L_b-1} ,
gdzie L_b jest liczbą bitów w reprezentacji próbek.

$F(k_1, k_2)$ – próbka transformaty,

$Q(k_1, k_2)$ – odpowiedni element tablicy kwantyzacji,

$Q_P(k_1, k_2)$ – przeskalowany element tablicy kwantyzacji,

$F_Q(k_1, k_2)$ – skwantowana próbka transformaty.

$$F_Q(k_1, k_2) = [F(k_1, k_2) / Q_P(k_1, k_2)] ,$$

gdzie $[\cdot]$ oznacza zaokrąglanie do wartości całkowitej.

Kwantowanie współczynników transformaty kosinusowej (c.d.)

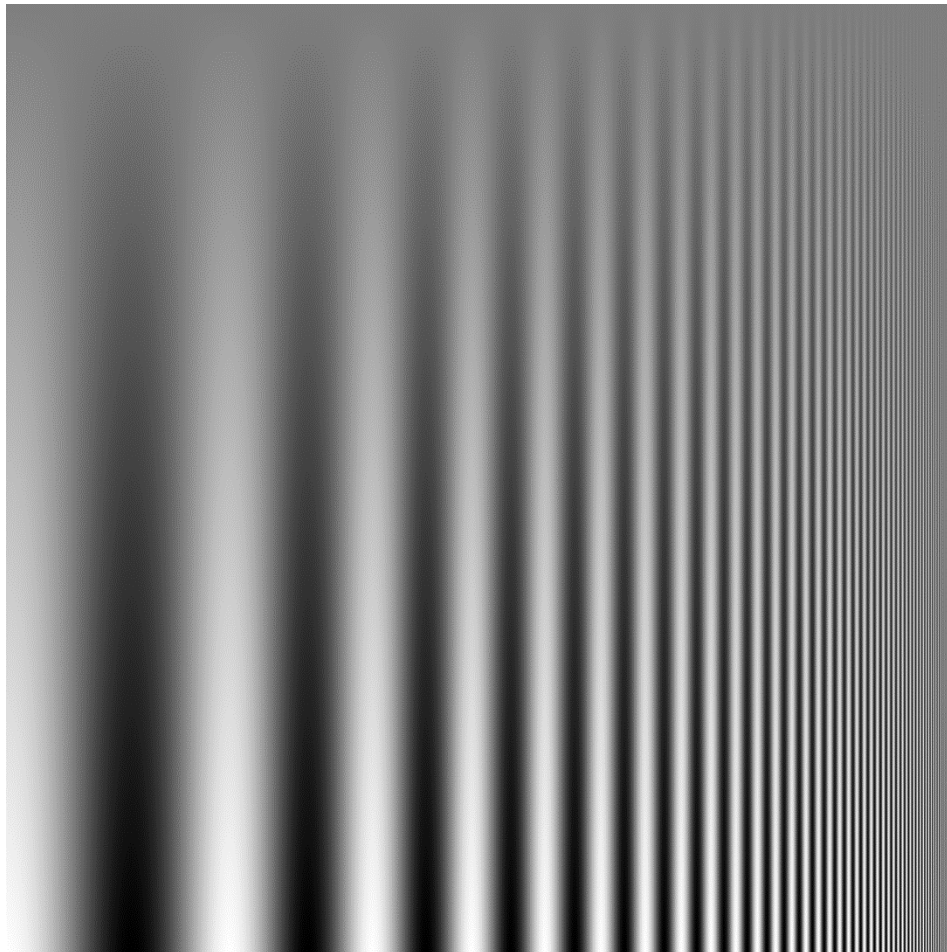
Podstawowa tablica kwantyzacji luminancji dla kodowania według normy JPEG

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Podstawowa tablica kwantyzacji chrominancji dla kodowania według normy JPEG

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Plansza Campbella-Robinsona.
Zmodulowane sinusoidalnie linie: częstotliwość
przestrzenna rośnie przy przesuwaniu się z lewa na prawo,
kontrast maleje przy przesuwaniu się ku górze.



Kwantowanie współczynników transformaty kosinusowej (c.d.)

Podstawowa tablica kwantyzacji luminancji dla kodowania według normy JPEG

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Podstawowa tablica kwantyzacji chrominancji dla kodowania według normy JPEG

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

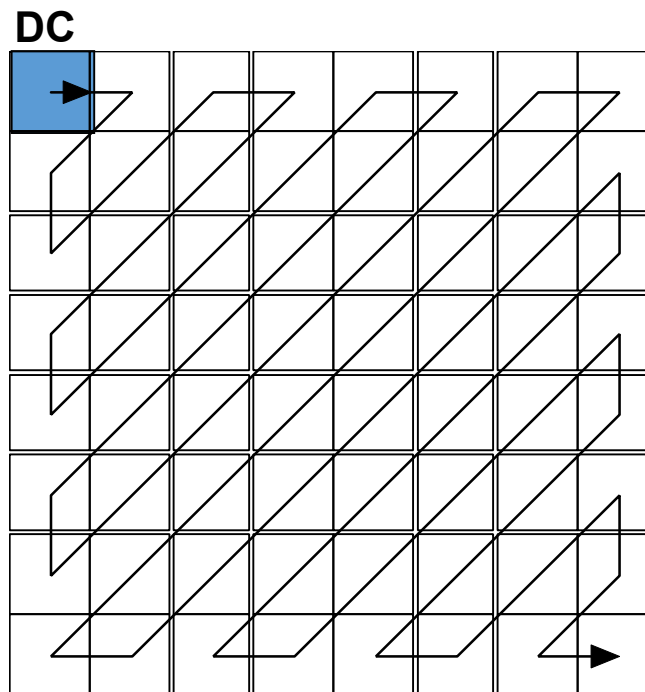
Kwantowanie współczynników transformaty kosinusowej

**Kwantowanie jest jedyną operacją
nieodwracalną.**

**Skalowanie tablicy kwantyzacji służy do regulacji
stopnia kompresji.**

7.6. Kodowanie skwantowanych współczynników transformaty kosinusowej

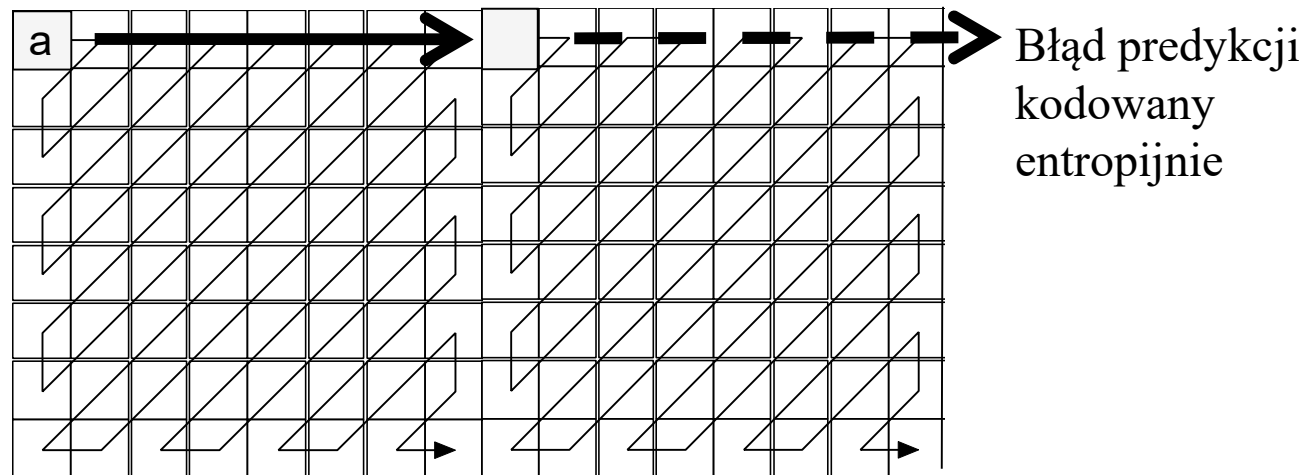
Uporządkowanie zygzakowate



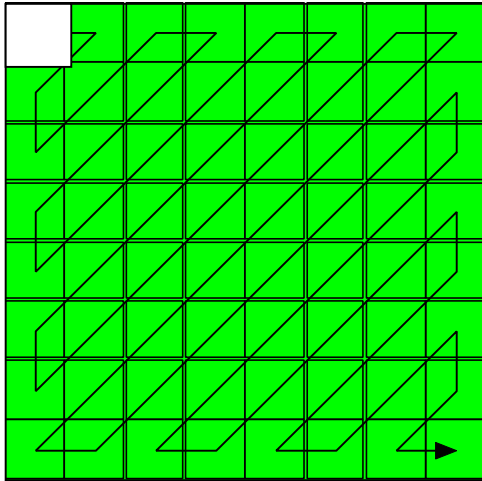
Kodowanie skwantowanych współczynników DC

- **Współczynniki DC** ($F(0, 0)$) transformaty kosinusowej dla poszczególnych bloków są **kodowane predykcyjnie**.
- Proces tego kodowania można sobie wyobrazić jako zastosowanie kodowania predykcyjnego do tablicy złożonej wyłącznie ze współczynników DC.
- Sygnał różnicowy jest kodowany entropijnie.

Predykcja współczynników DC



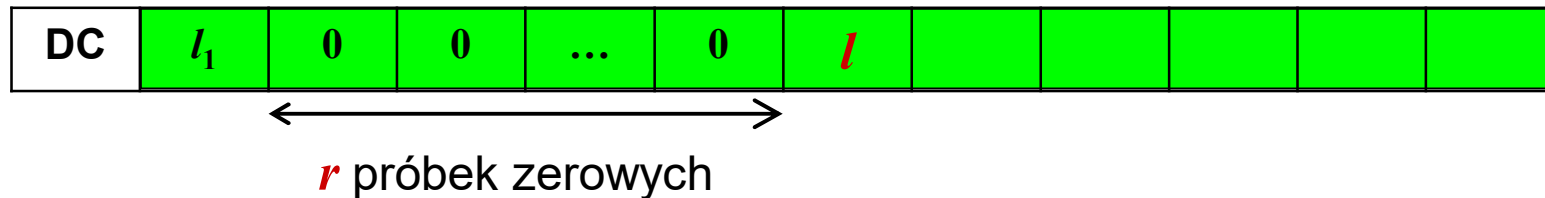
Kodowanie współczynników AC



Współczynniki AC są kodowane blokami.

Próbki widma są porządkowane liniowo wewnątrz poszczególnych bloków.

Kodowanie ciągów próbek widma (współczynników transformaty):



Współczynniki AC bloku są reprezentowane ciągiem:

$(r_1, l_1), (r_2, l_2), (r_3, l_3), \dots, \text{EOB}$ (EOB – *end of block* - koniec bloku)

Kodowanie skwantowanych współczynników AC (c.d.)

Wartość l jest przedstawiana jako

$l = z \cdot 2^{(s-1)} + a$, gdzie $z = \pm 1$, a liczba s określa zakres wartości l .

Za pomocą kodu Huffmana przedstawia się pary (r, s) i dodatkowo przesyła się wartości a i z .

Wartości zmiennej s dla kodowania współczynników AC w trybie sekwencyjnym przy 8-bitowych danych wejściowych

Zakres wartości współczynnika AC	s
-1, 1	1
-3, -2, 2, 3	2
-7,..., -4, 4,..., 7	3
-15,..., -8, 8,..., 15	4
-31,..., -16, 16,..., 31	5
-63,..., -32, 32,..., 63	6
-127,..., -64, 64,..., 127	7
-255,..., -128, 128,..., 255	8
-511,..., -256, 256,..., 511	9
-1023,..., -512, 512,..., 1023	10

Kodowanie Huffmana współczynników AC

Fragment tablicy podstawowych kodów Huffmana dla współczynników AC lumpy dla liczby powtórzeń r nie większej od 2

r/s	Kod	r/s	Kod
0/1	00	1/6	1111111110000100
0/2	01	1/7	1111111110000101
0/3	100	1/8	1111111110000110
0/4	1011	1/9	1111111110000111
0/5	11010	1/10	1111111110001000
0/6	1111000	2/1	11100
0/7	11111000	2/2	11111001
0/8	1111110110	2/3	1111110111
0/9	1111111110000010	2/4	111111110100
0/10	1111111110000011	2/5	1111111110001001
1/1	1100	2/6	1111111110001010
1/2	11011	2/7	1111111110001011
1/3	1111001	2/8	1111111110001100
1/4	111110110	2/9	1111111110001101
1/5	11111110110	2/10	1111111110001110

7.7. Norma JPEG

Podstawowa norma:

ISO/IEC IS 10918-1 / ITU-T Rec. T.81 (ustalona w roku 1992)

Information technology – digital compression and coding of continuous-tone still images: requirements and guideline.

Uzupełnienia: (nie należy się tego uczyć na pamięć)

- ISO/IEC IS 10918-2 / ITU-T Rec. T.83

Information technology – digital compression and coding of continuous-tone still images: compliance testing.

- ISO/IEC IS 10918-3 / ITU-T Rec. T.84

Information technology – digital compression and coding of continuous-tone still images: extensions.

- ISO/IEC IS 10918-4 / ITU-T Rec. T.86

Information technology – digital compression and coding of continuous-tone still images: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF colour spaces, APPn markers, SPIFF compression types and Registration Authorities (REGAUT)

- ISO/IEC IS 10918-5 / ITU-T Rec. T.871

Information technology – digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF)

- ISO/IEC IS 10918-6 / ITU-T Rec. T.872

Information technology – digital compression and coding of continuous-tone still images: Application to printing systems

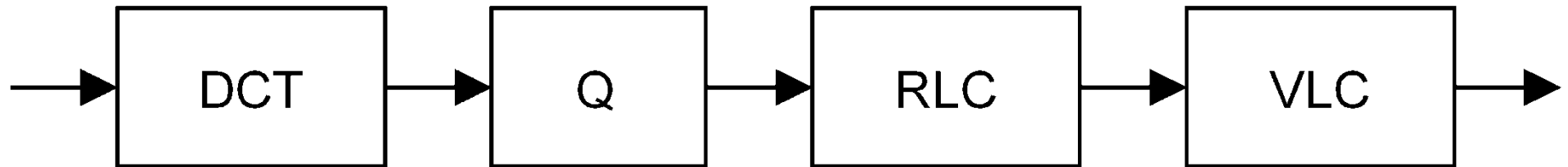
- ITU-T Rec. T.851

ITU-T T.81 (JPEG-1)-based still-image coding using an alternative arithmetic coder

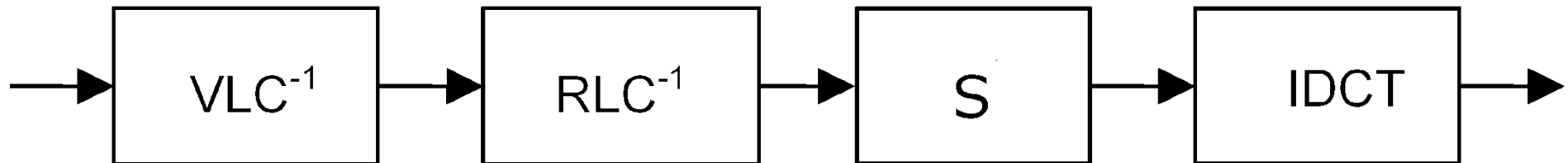
Kodek JPEG

- Najczęściej używany kodek dla kompresji stratnej obrazów nieruchomych – wiele miliardów urządzeń oraz egzemplarzy oprogramowania.
- Prosty, w miarę efektywny.
Małe zużycie energii w urządzeniach przenośnych.
- Wolny od patentów.
- Stara technika (1992 – norma).

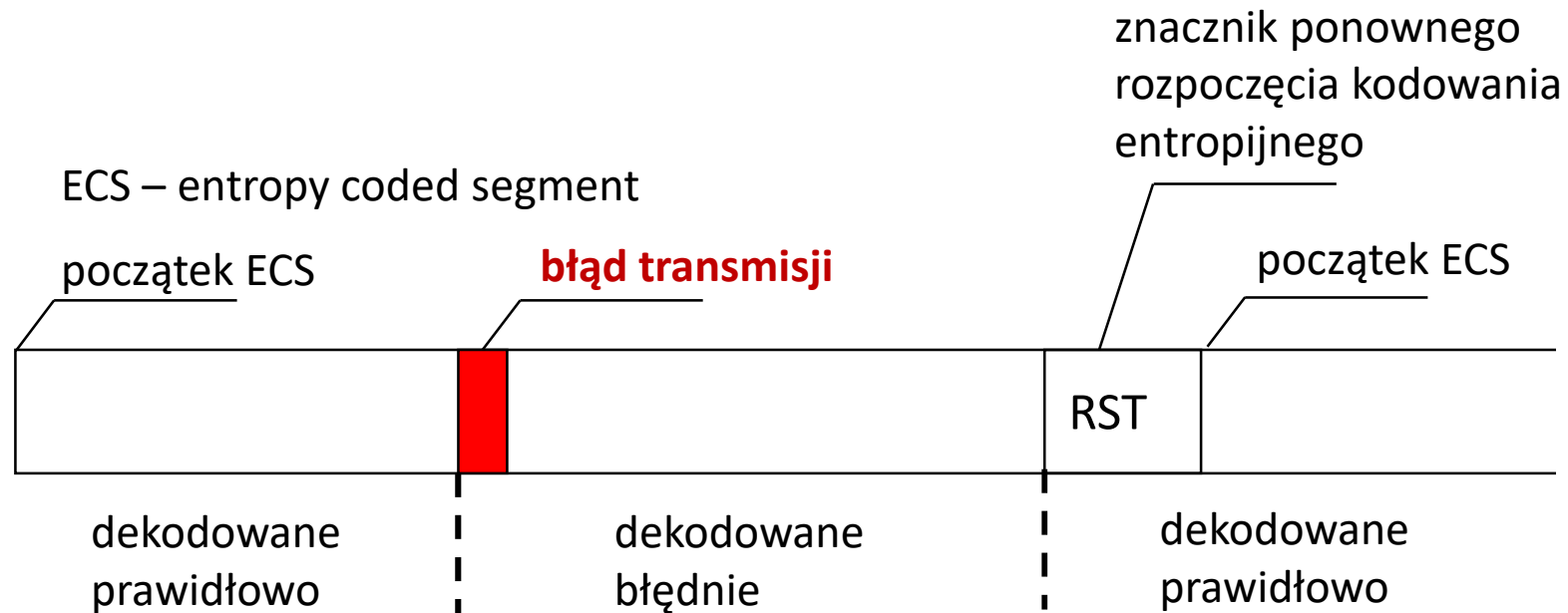
Koder



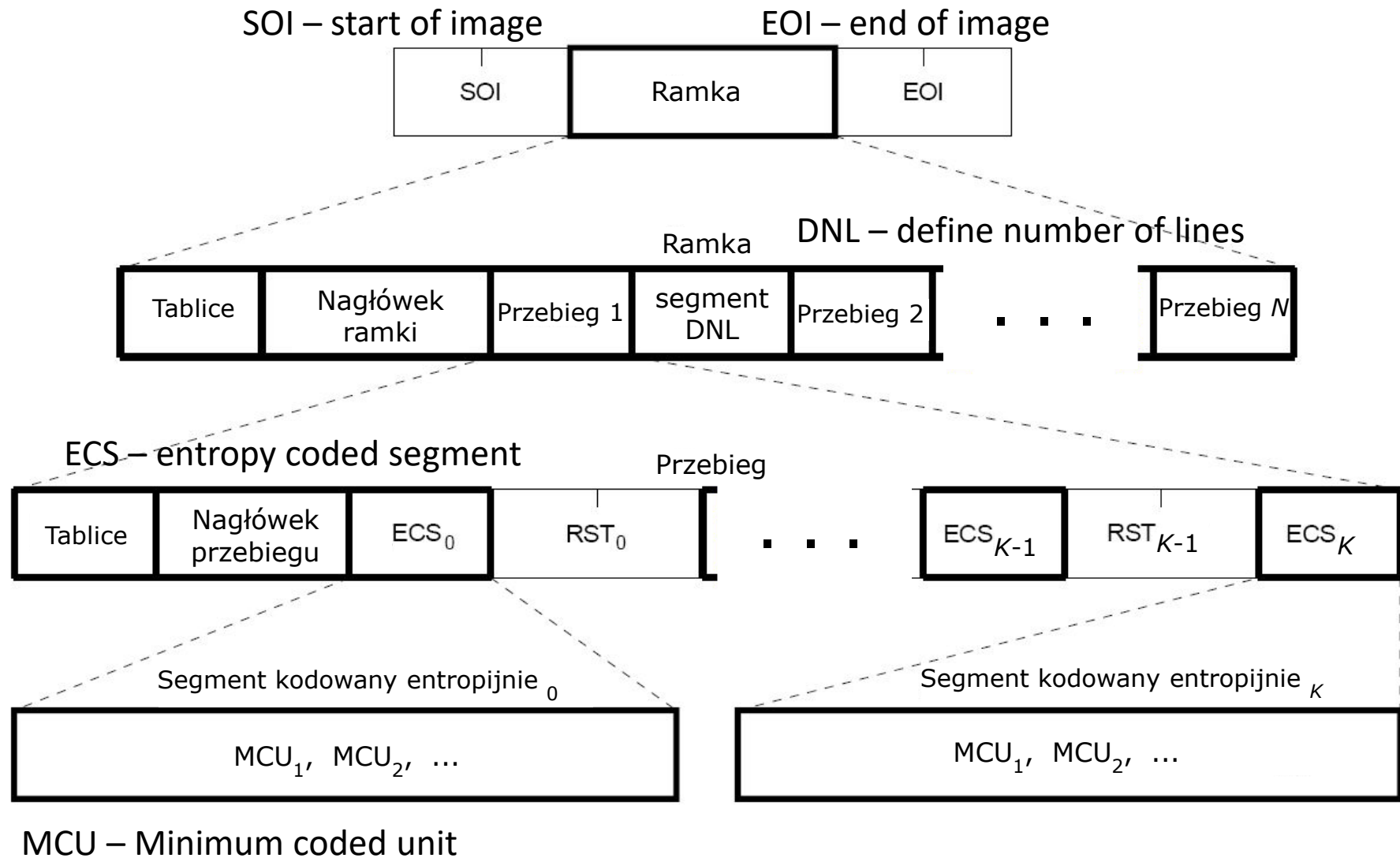
Dekoder



Redukcja wpływu błędów transmisji na jakość obrazu zdekodowanego



Strumień danych



Uzyskiwana kompresja:

Efektywność kodera JPEG w podstawowym trybie sekwencyjnym:
Dla składowej lummy:

- **Oryginał jest nierozróżnialny** od obrazu zdekodowanego dla strumienia binarnego o wartościach **1,5 – 2,0 b/punkt**.
- **Bardzo dobrą lub dobrą jakość** uzyskuje się w zakresie od około **0,5 do 1,5 b/punkt**.
- Wartości współczynnika kompresji **przekraczające 40** są związane z kiepską lub złą jakością obrazu.
- Zastąpienie standardowych tablic kwantyzacji oraz kodów VLC tablicami dostosowanymi do statystyki kodowanego obrazu może poprawić wartość PSNR o około 1 – 2 dB.

Efekt blokowy

- Efekt silnej kompresji.
- Pojawia się na granicach bloków 8x8.
- Na obrazku sztucznie wzmocniony.



Przykłady – obraz oryginalny (RGB) – 768 kB (24b/p)



Przykłady – obraz odtworzony (RGB) – 53 kB (1,6b/p) $k = 15$



Przykłady – obraz odtworzony – 22,2 kB (0,67b/p) $k = 36$



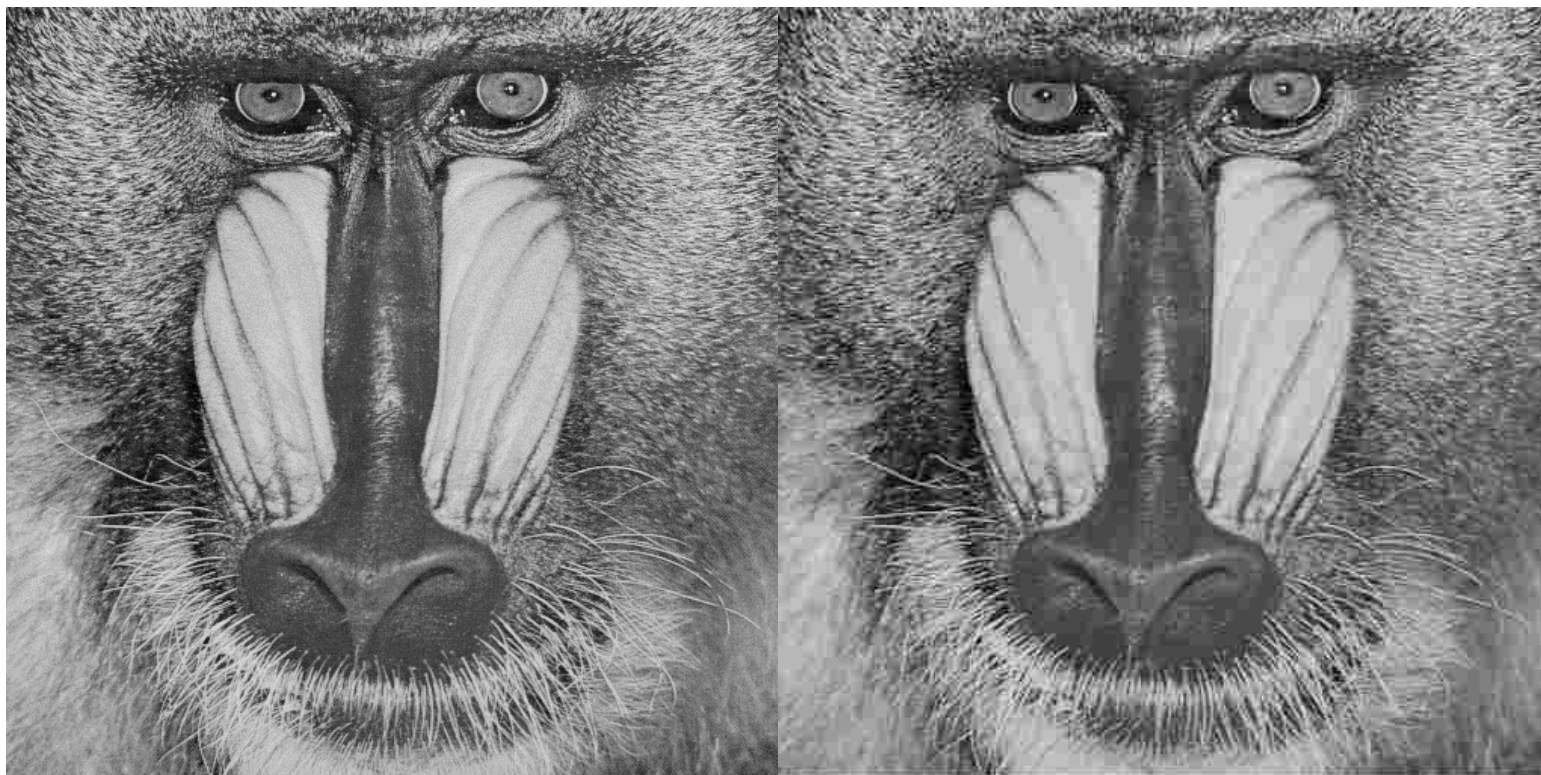
Oryginał

Odtworzony 0,5 b/p

512x512 Lena



512x512 baboon



Formaty plików

- JIF – JPEG Interchange Format,
- JFIF – JPEG File Interchange Format ,
- EXIF – Exch.angeable Image File Format

7.8. Kodowanie skalowalne

Obraz kodowany w postaci kilku warstw.

Warstwa podstawowa – część reprezentacji bitowej umożliwiające zdekodowanie obrazu o słabej jakości.

Warstwa uzupełniająca - część reprezentacji bitowej, której zdekodowanie umożliwia poprawę jakości.

Skalowalność jakości – skalowalność SNR

Tryb progresywny kodowania transformatowego.

Składowe obrazu są kodowane przez wieloetapowe przetwarzanie.

W każdym etapie kodowana jest tylko część informacji o skwantowanych współczynnikach DCT.

Tryb jest użyteczny dla powolnych transmisji.

Umożliwia odbiorcy szybkie uzyskanie wersji o gorszej jakości, która w kolejnych etapach ulega poprawie aż do osiągnięcia pełnej jakości.

Są stosowane dwie metody:

- selekcja widmowa (*spectral selection*),
- kolejne aproksymacje (*successive approximations*).

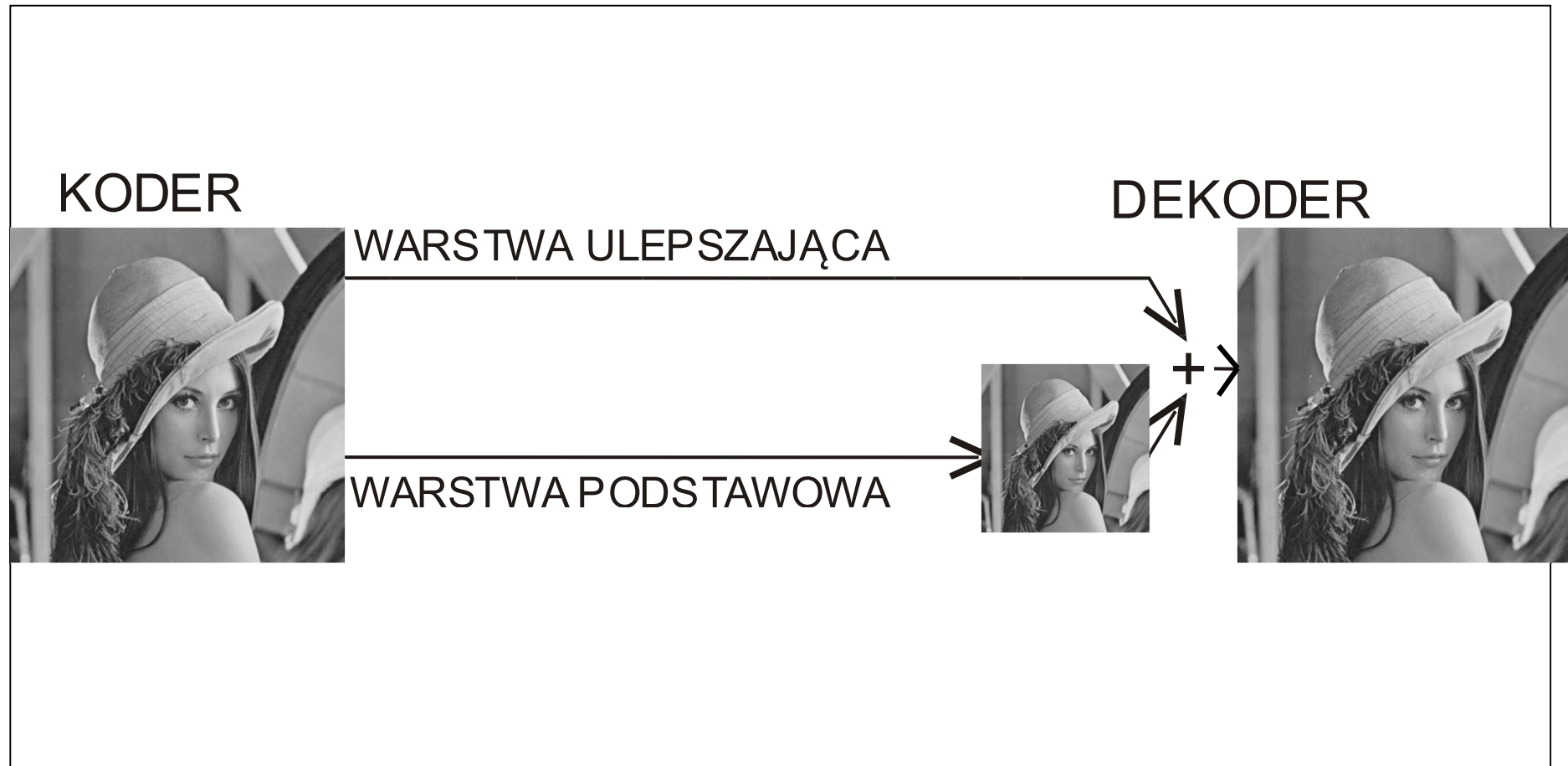
Określa się 4 tryby pracy (c.d.):

Tryb **bezstratny**

Tryb **hierarchiczny**

określa kolejne kodowanie obrazów o rosnącej rozdzielczości. Pozwala na transmisję i zdekodowanie obrazu o małej rozdzielczości bez potrzeby transmisji i dekodowania obrazu o pełnej rozdzielczości. W dekodерze z obrazu oryginalnego poprzez kolejne decymacje są wytwarzane obrazy o coraz mniejszej rozdzielczości.

Tryb hierarchiczny



7.9. Porównania technik kompresji stratnej

Podstawowe techniki kompresji stratnej pojedynczych obrazów

- **JPEG** ISO/IEC IS 10918, ITU-T Rec. T.81 itd.
Kodowanie blokowe transformatowe, tablice kwantyzacji, kodowanie Huffmana (ew. arytmetyczne) .
- **JPEG 2000** ISO/IEC IS 15444, ITU-T Rec. T.800 itd.
Kodowanie falkowe, złożona alokacja bitów (EBCOT –*Embedded Block Coding by Optimized Truncation*), kodowanie arytmetyczne.
Kompresja stratna i bezstratna.
- **JPEG XR** ISO/IEC IS 29199, ITU-T Rec. T.832 itd.
Dwa etapy kodowania transformatowego w blokach nakładających się, kwantowanie modyfikowane w poszczególnych blokach, kodowanie o zmiennej długości słowa.
Bazuje na *Microsoft HD Photo*
- **WebP** RFC 9649 (Internet Engineering Task Force – IETF)
Predykcja w blokach. Transformacje kosinusowa i Walsha-Hadamarda.
Technika kompresji stratnej i bezstratnej.

Porównanie

- **JPEG a JPEG 2000**

JPEG 2000: zwykle plik mniejszy o 15-40%
Czasy kodowania i dekodowania dłuższe ok. 5-10 razy.

- **JPEG a JPEG XR**

JPEG XR: zwykle plik mniejszy o 10-30%
Czasy kodowania i dekodowania nieco dłuższe, ale znacznie krótsze niż dla JPEG2000.

Porównanie efektywności kodowania

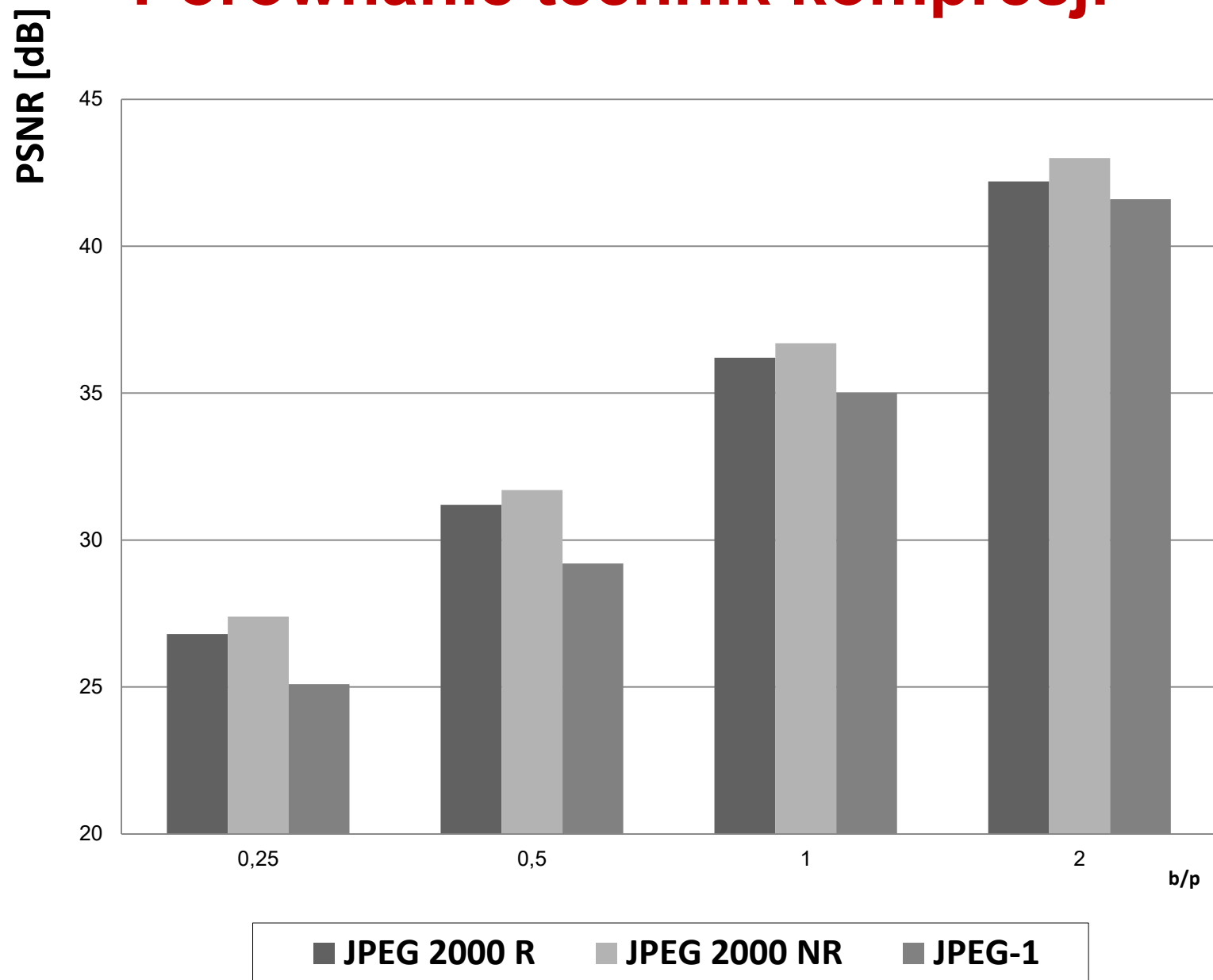
JPEG-1 at 0.27 bpp

Original

JPEG-2000 at 0.27 bpp



Porównanie technik kompresji



Porównanie JPEG-1 i JPEG 2000

Oryginał



JPEG
0,1 b/p

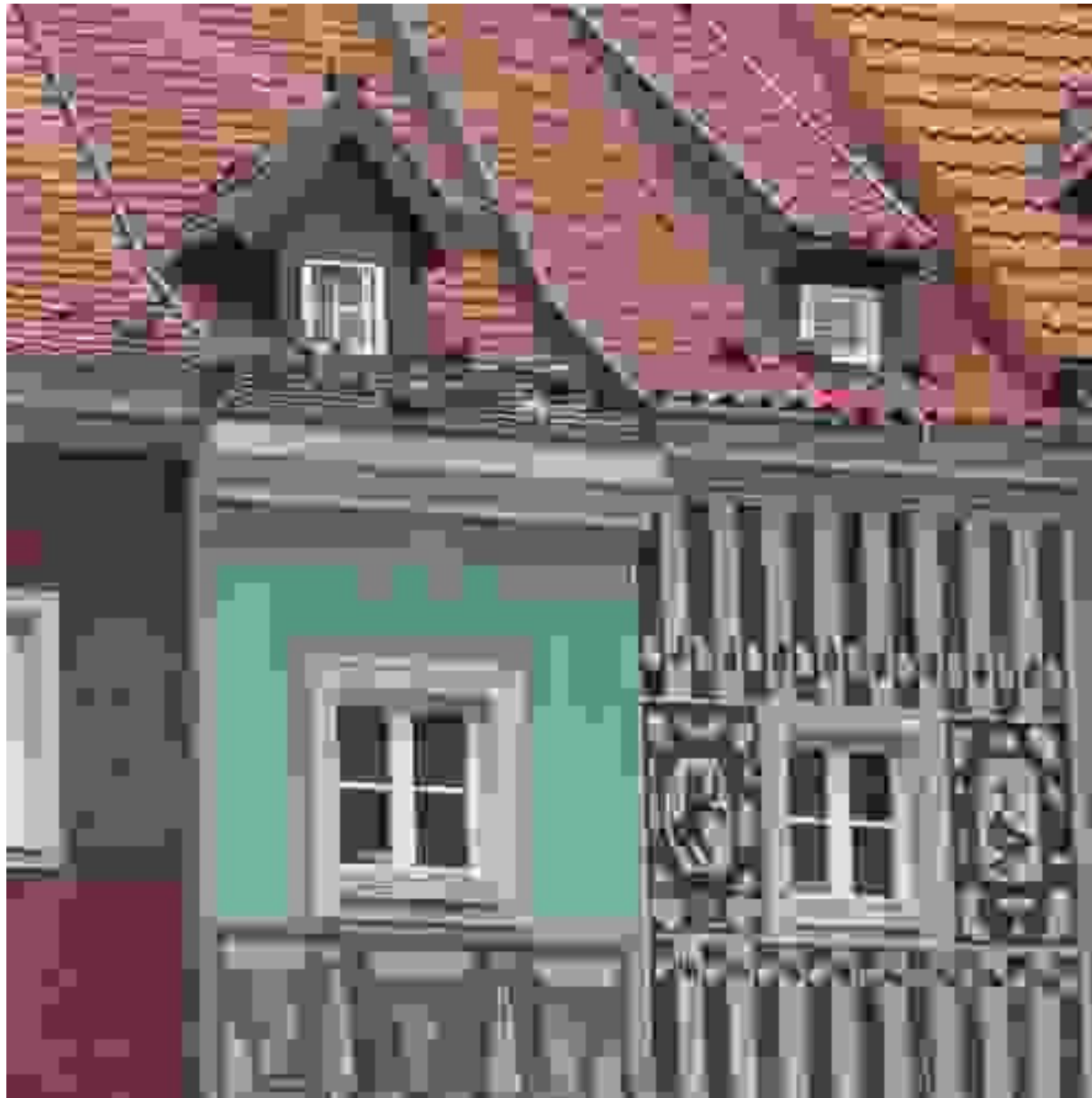


JPEG 2000
0,1 b/p





Oryginał



JPEG
(0,2 b/p)

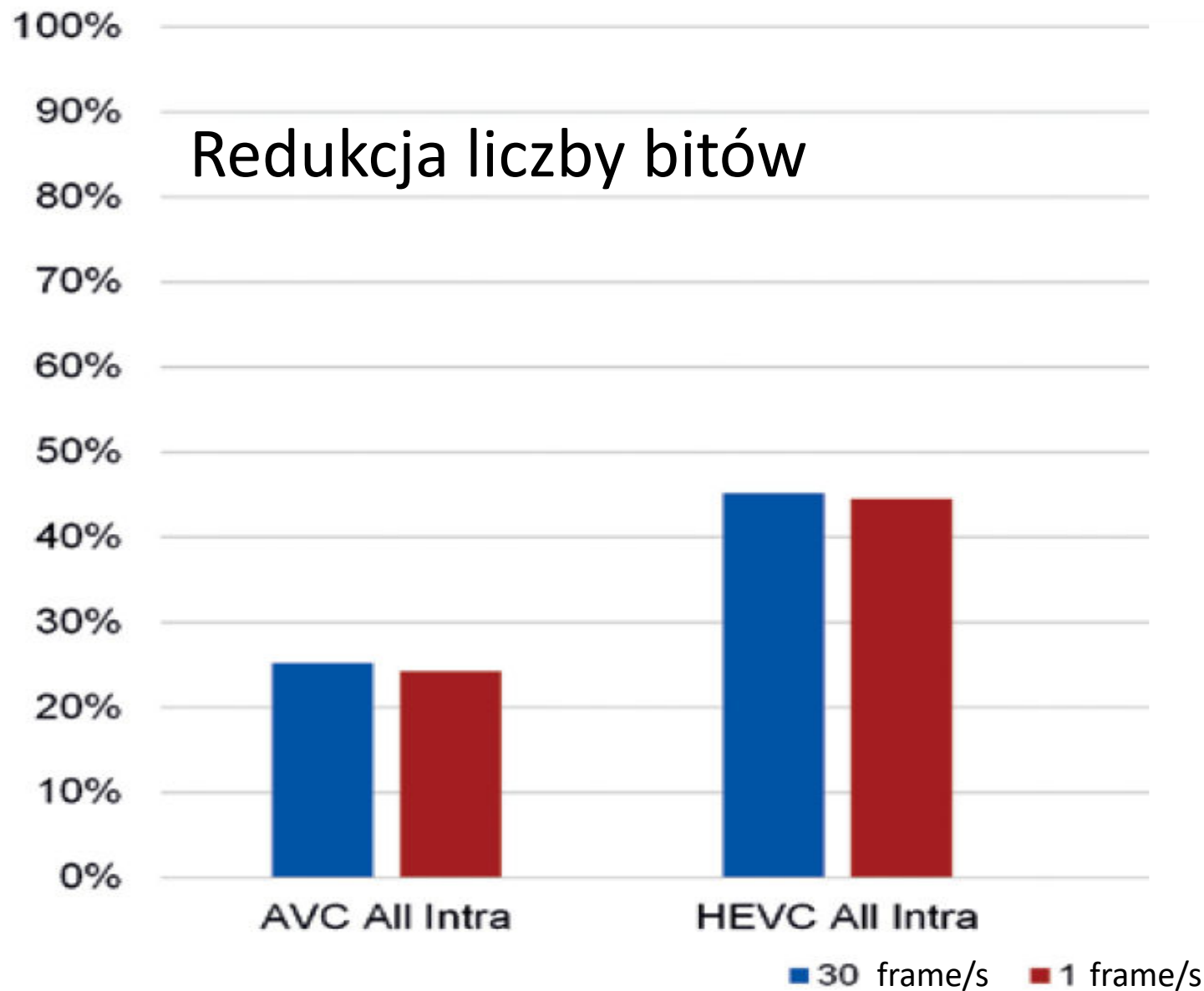


JPEG2000
(0,2 b/p)

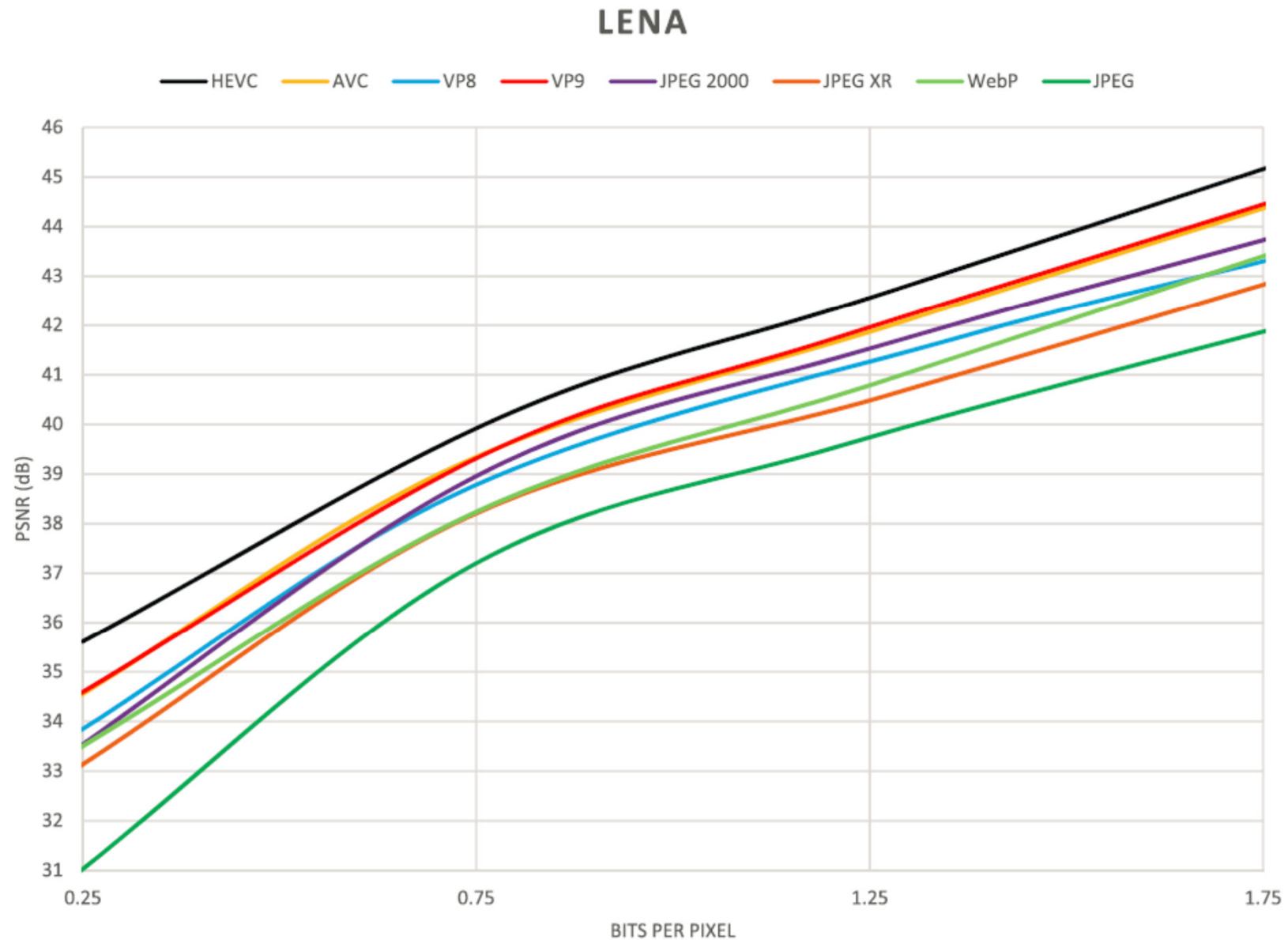
Kodowanie wewnątrzbrazowe w technikach kompresji wizji

- **AVC** – Advanced Video Coding
ISO/IEC IS 14496-10 = MPEG-4 Part 10
ITU-T Rec. H.264.
- **HEVC** – High Efficiency Video Coding
ISO/IEC IS 23008-2 = MPEG-H Part 2
ITU-T Rec. H.265.
- **VVC** – Versatile Video Coding
ISO/IEC IS 23090-3 = MPEG-I Part 3
ITU-T Rec. H.266.

Porównanie w stosunku do JPEG 2000

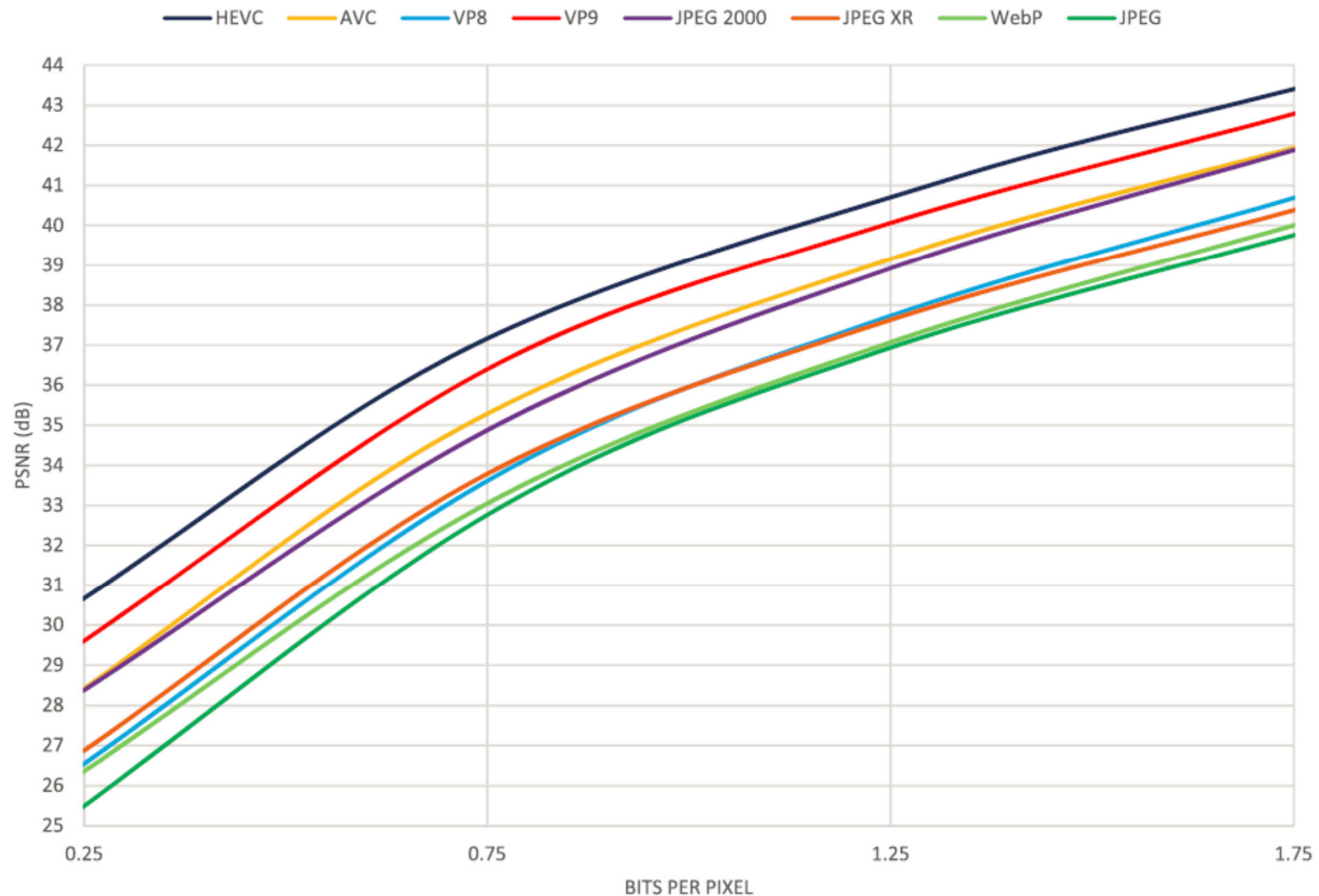


Krzywe R-D dla różnych kodeków



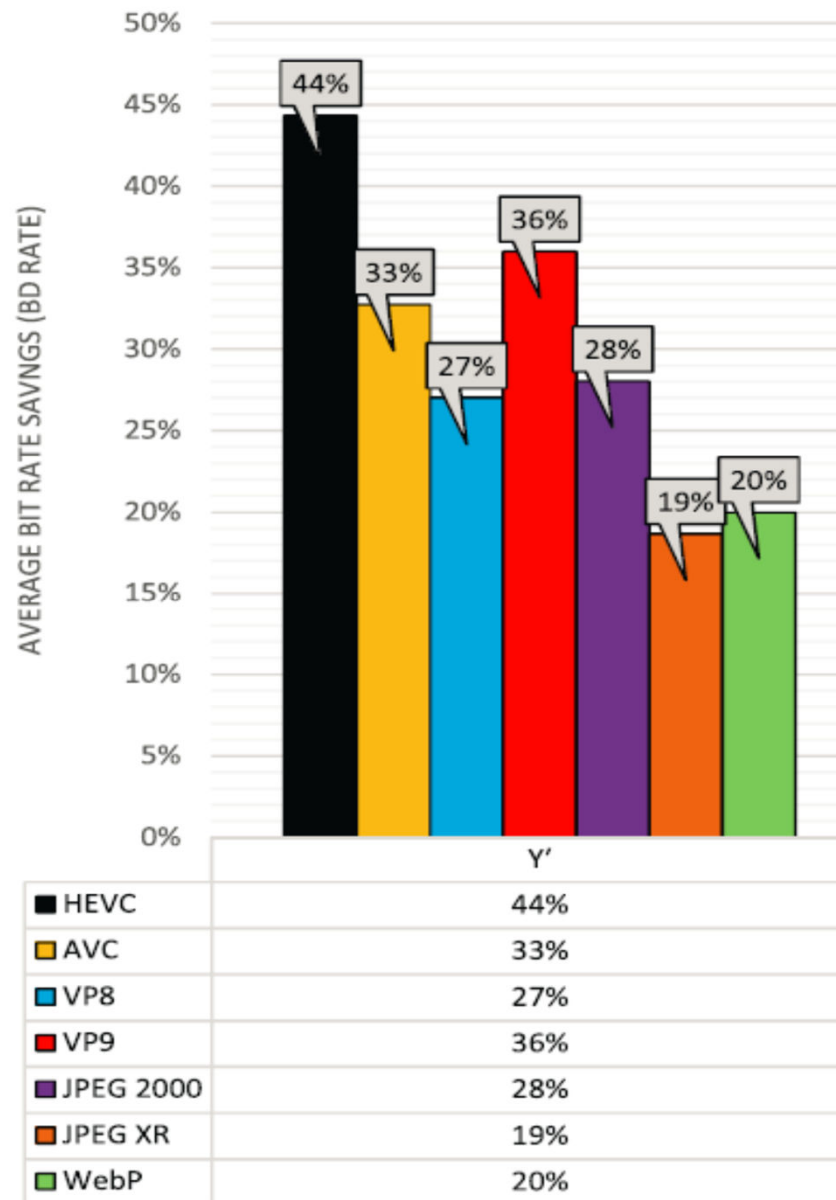
Krzywe R-D dla różnych kodeków

BARBARA



From: NGUYEN AND MARPE: OBJECTIVE PERFORMANCE EVALUATION OF THE HEVC MSP PROFILE

Porównanie w stosunku do JPEG



Zbiór testowy :
32 obrazy (JPEG/Kodak).
Jakość: PSNR

Marek Domański

Materiały pomocnicze do wykładu

Wprowadzenie do multimediiów

Politechnika Poznańska

© Marek Domański

Materiały są przeznaczone wyłącznie do użytku podczas zajęć z przedmiotu *Wprowadzenie do multimediiów* prowadzonych przez prof. Marka Domańskiego na studiach na Wydziale Elektroniki i Telekomunikacji.

Uwaga: Niektóre ilustracje pochodzą z podręczników zalecanych jako pomocnicze do przedmiotu.

Kopiowanie, rozpowszechnianie i używanie w innych celach jest stanowczo zabronione.

W szczególności, nieuprawnione rozpowszechnianie drogą elektroniczną np. w Internecie jest nielegalne

8. Kodowanie wizji

8.1. Obecnie używane kodeki wizji MPEG

1. Zaawansowane kodowanie wizji

Advanced Video Coding – AVC

Norma:

ISO/IEC IS 14496-10

**Information Technology – Coding of audio-visual
objects**

Part 10: Advanced Video Coding

także :

ITU-T Rec. H.264

Techniki kompresji ujęte w normie MPEG-4 czyli ISO/IEC IS 14496

- 14496-2: (MPEG-4)
- 14496-10: AVC (= ITU-T Rec. H.264)
- 14496-31: Video Coding for Browsers
- 14496-33: Internet Video Coding

2. Bardzo sprawne kodowanie wizji High Efficiency Video Coding – HEVC

- **ISO/IEC 23008-2 = MPEG-H Part 2**
- **ITU-T Rec. H.265**

Podstawa wykładu

3. Wszechstronne kodowanie wizji Versatile Video Coding – VVC

- **ISO/IEC 23090-3 = MPEG-I Part 3**
- **ITU-T Rec. H.266**

4. Zasadnicze kodowanie wizji

Essential Video Coding – EVC

- **ISO/IEC 23094-1 = MPEG-5 Part 1**
Video coding format
Podstawowy profil ma być wolny od opłat patentowych (royalty-free).
- **ISO/IEC 23094-2 = MPEG-5 Part 2**
Low Complexity Enhancement Video Coding

Nie będzie rozpatrywany na wykładzie.

Poza ISO/IEC (MPEG) i ITU: Inne współczesne kodeki o istotnym znaczeniu

❖ AV1

- ◆ Konsorcjum firm: **Alliance for Open Media**: Amazon, Apple, ARM, Cisco, Facebook, Google, IBM, Intel, Microsoft, Mozilla, Netflix, Nvidia, Samsung Electronics and Tencent i inne.
- ◆ Specyfikacja opublikowana w 2018.
- ◆ Unikanie patentów nie będących własnością konsorcjum.
- ◆ Stopień kompresji podobny do HEVC.
- ◆ Implementowane w Windows (2020).

❖ AVS Audio Video Coding Standard of China

- ◆ Wersja 1 (2006): odpowiednik AVC.
- ◆ Wersja 2 (2016): częściowy odpowiednik HEVC.
- ◆ Wersja 3 (2019): wykorzystuje wiedzę zdobytą przy opracowywaniu VVC.

8.2. Profile, warstwy i poziomy kodeków

- **Profil (*profile*)**

- ◆ Podzbiór zaleceń normy (podzbiór składni strumienia) dla pewnego obszaru zastosowań.

- ◆ Na przykład:

- profil dla kodowania wizji dla telewizji i filmów w internecie,

- profil dla kodowania wizji stereoskopowej,

- profil dla kodowania wizji monochromatycznej,

- profil dla kodowania wizji ze schematem próbkowania chromy 4:2:2,

- profil dla kodowania składowych o 12 bitach na próbkę,

- itp.

Profile, warstwy i poziomy kodeków

- **Warstwa (*tier*)**

Określa odmianę poziomu: np.

1) dla zastosowań typowych dystrybucyjnych (np. telewizja, media w IP), w tym konsumenckich.

2) dla zastosowań szczególnie wymagających, np. w środowisku produkcyjnym. (brak w AVC).

Profile, warstwy i poziomy kodeków

- **Poziom (*level*)**
- Dopuszczalne zakresy wartości parametrów ilościowych obrazu i strumienia dla kodeka o określonym profilu.
Np. maksymalna liczba próbek lumy w obrazie, maksymalna prędkość bitowa po kompresji, maksymalna liczba próbek obrazu na sekundę itp.
- Ograniczenia liczbowe są tak zdefiniowane dla poszczególnych profili, że poszczególne poziomy odpowiadają pewnym formatom wizji.

Przykłady:

SDTV (576 linii, 25/30 obraz/s)	3
HDTV (1080 linii, 25/30 obraz/s)	4
HDTV (1080 linii, 50/60 obraz/s)	4.1
UHDTV (2160 linii, 25/30 obraz/s)	5
UHDTV (2160 linii, 50/60 obraz/s)	5.1

8.3. Podstawowe narzędzia

8.3.1. Estymacja ruchu

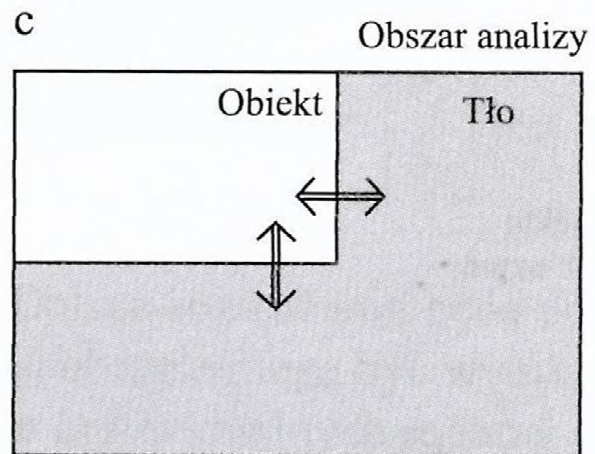
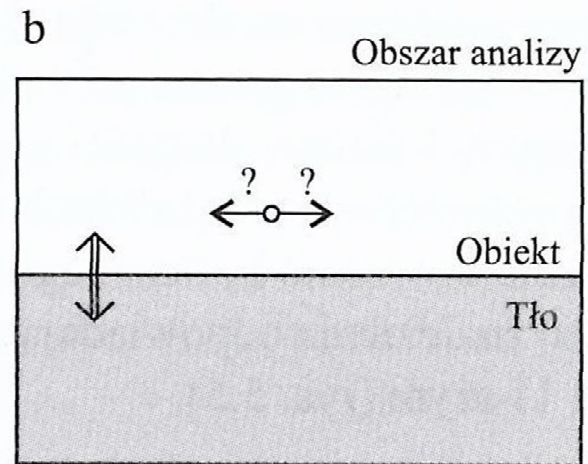
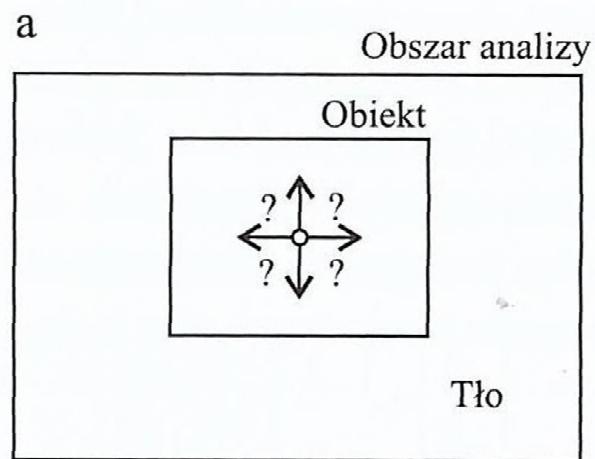
Cel: wyznaczenie wektorów ruchu

Pole ruchu:

- **gęste** — metody przepływu optycznego
- **rzadkie** — metody pasowania bloków

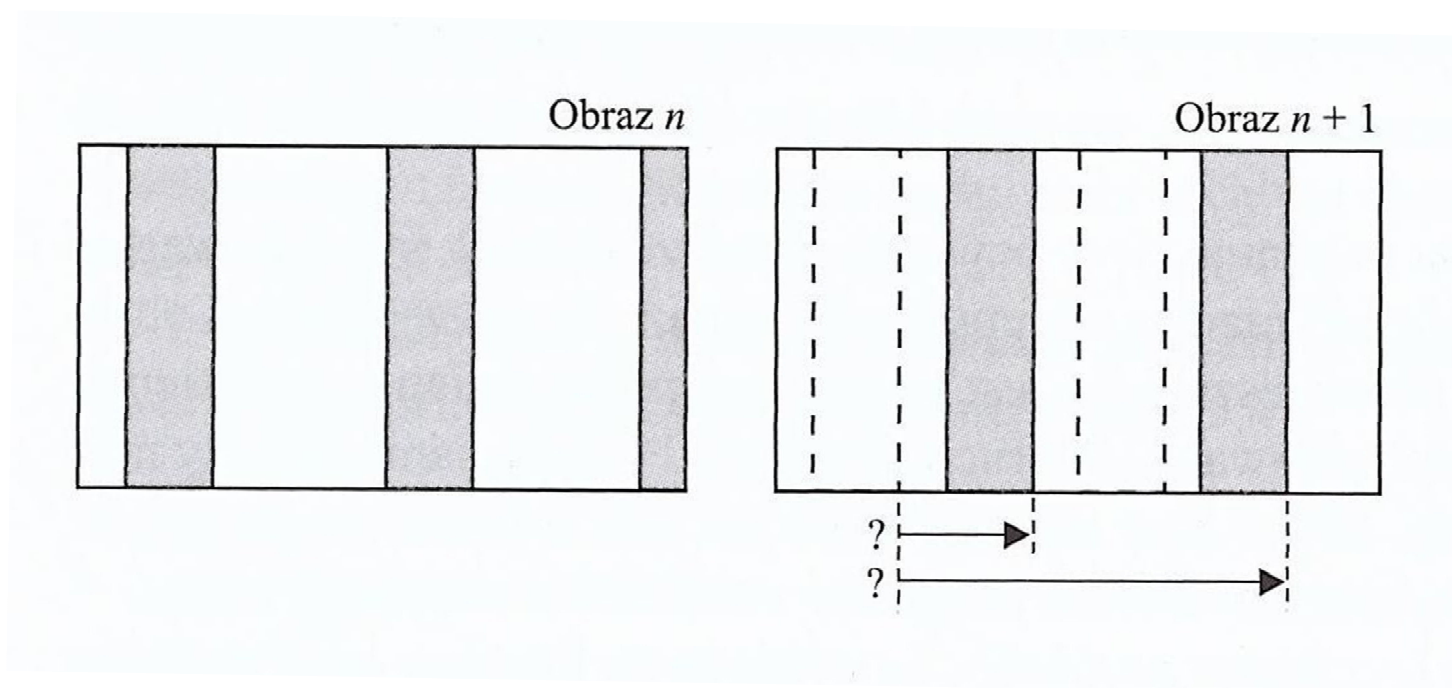
Problemy

Problem aperture

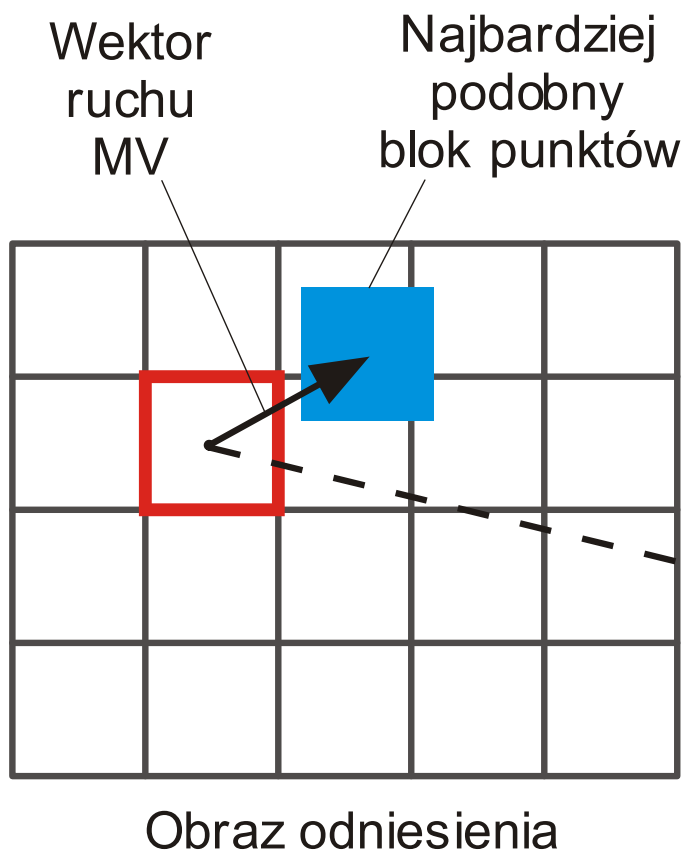


Problem przesłaniania

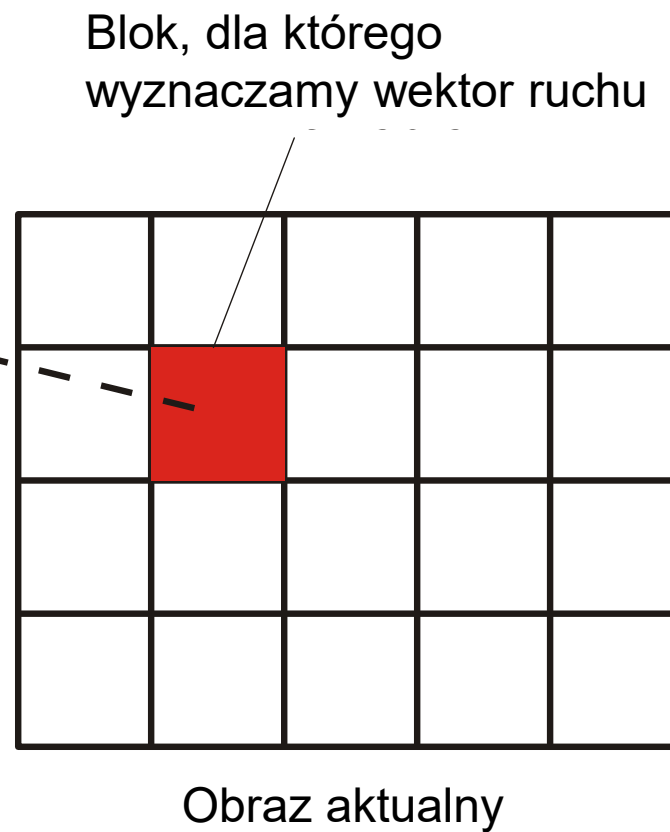
Problem odpowiedniości



Estymacja ruchu

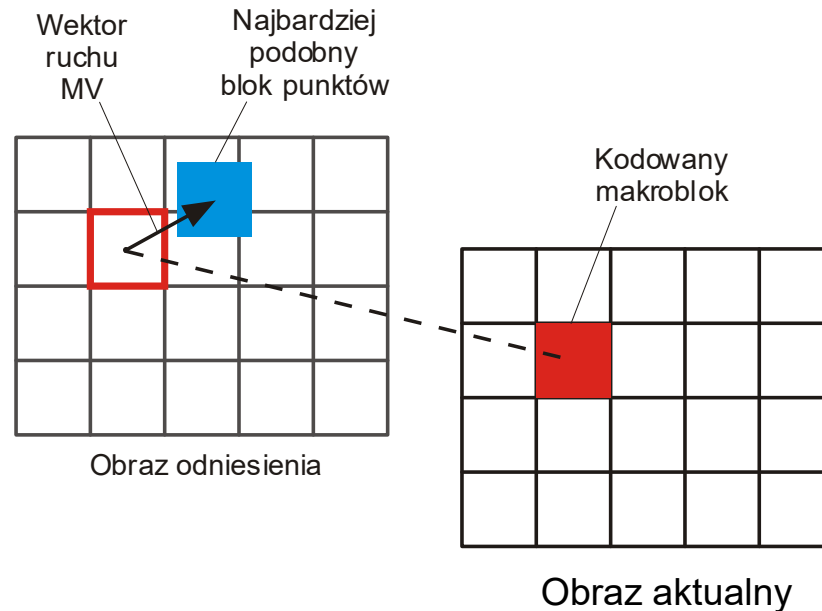


Wektory ruchu wyznacza się dla składowej luminancji (tylko !)



Estymacja ruchu (c.d.) - **Metoda pasowania bloków**

Poszukiwana jest para liczb (d_1, d_2) **minimalizująca** wskaźnik



$$E(n_1, n_2, n_t; d_1, d_2) =$$
$$= \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F[u_{akt}(n_1 + k_1, n_2 + k_2, n_t), u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)],$$

Najczęściej stosowane normy – miary błędu :

- kryterium MAE (*mean absolute error*) lub MAD (*mean absolute difference*)

$$F(u_1, u_2) = |u_1 - u_2|,$$

- kryterium MSE (*mean square error*)

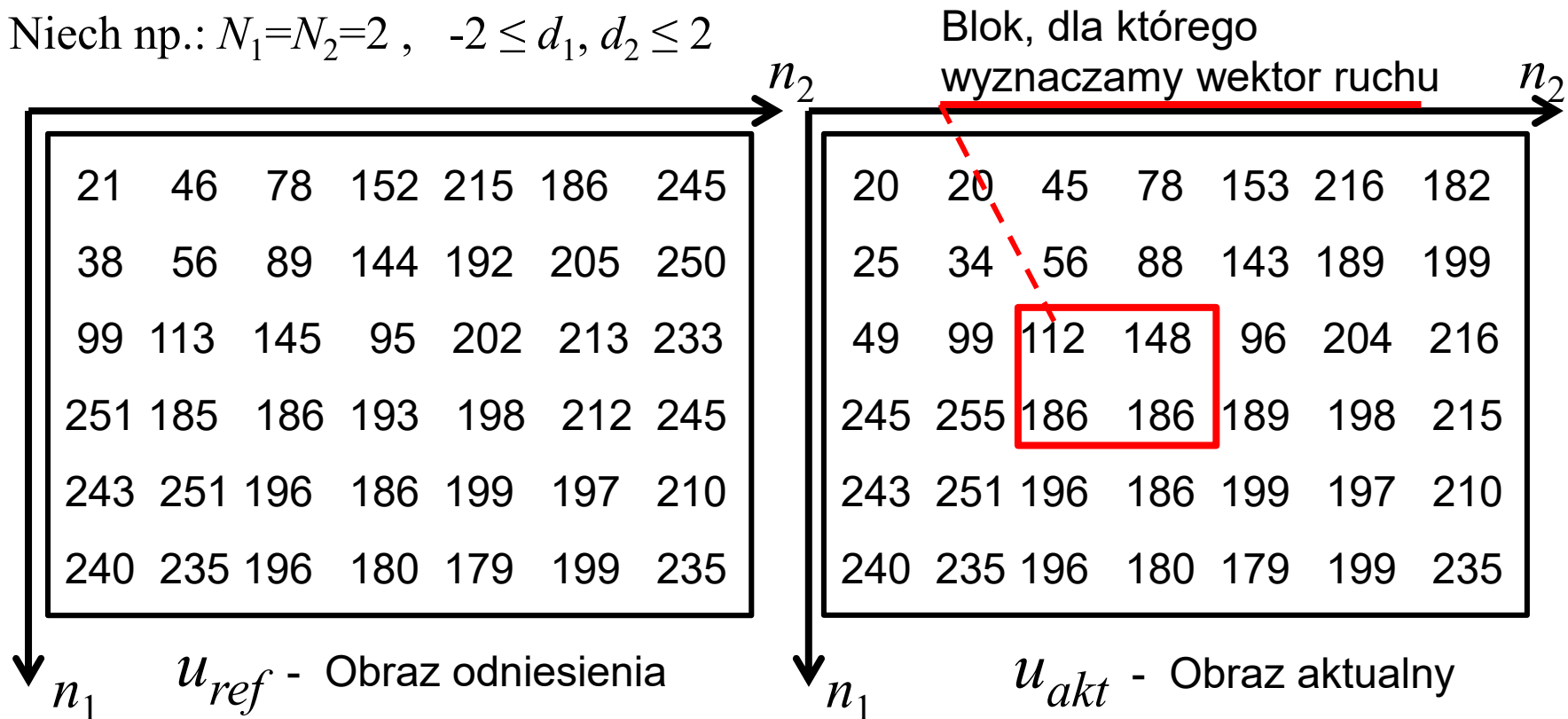
$$F(u_1, u_2) = (u_1 - u_2)^2.$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) =$$

$$= \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

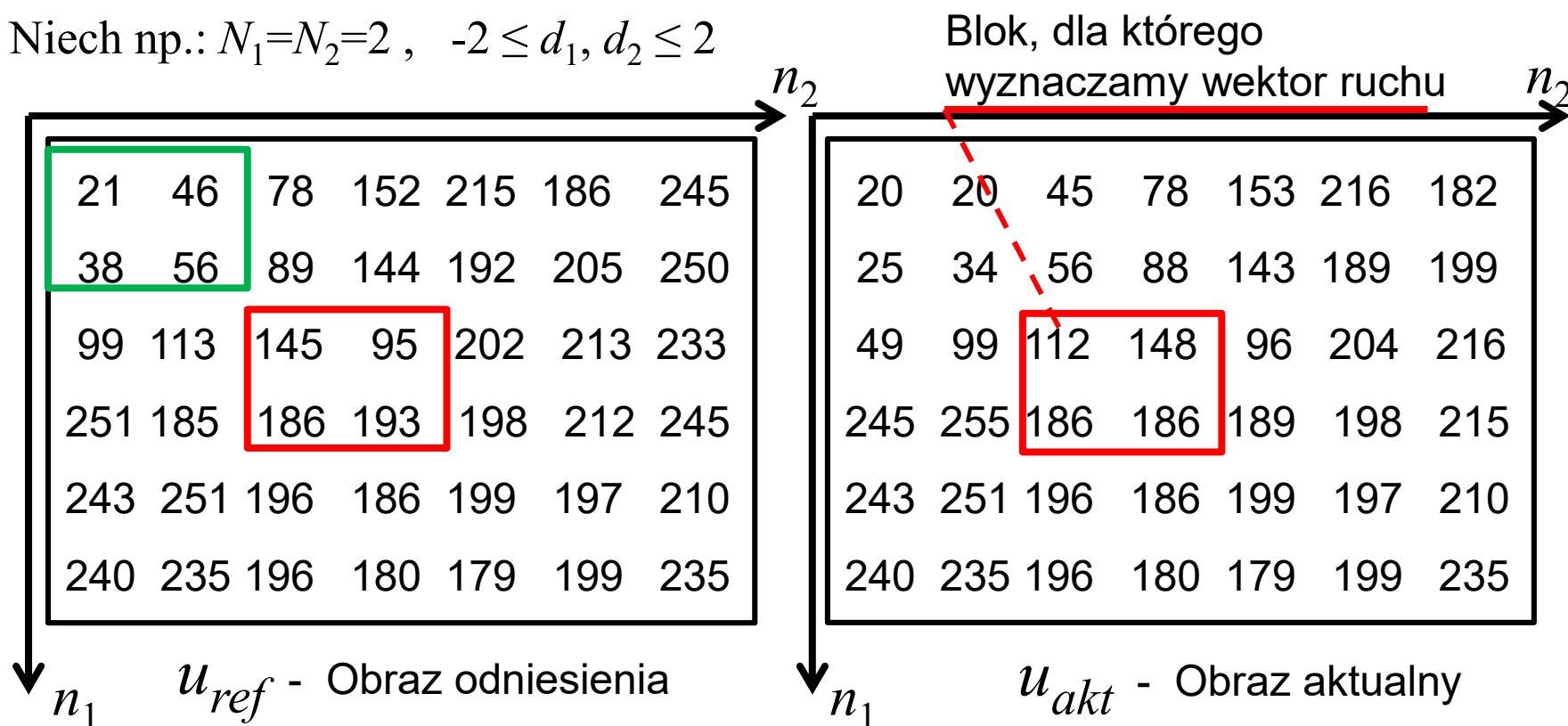
Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



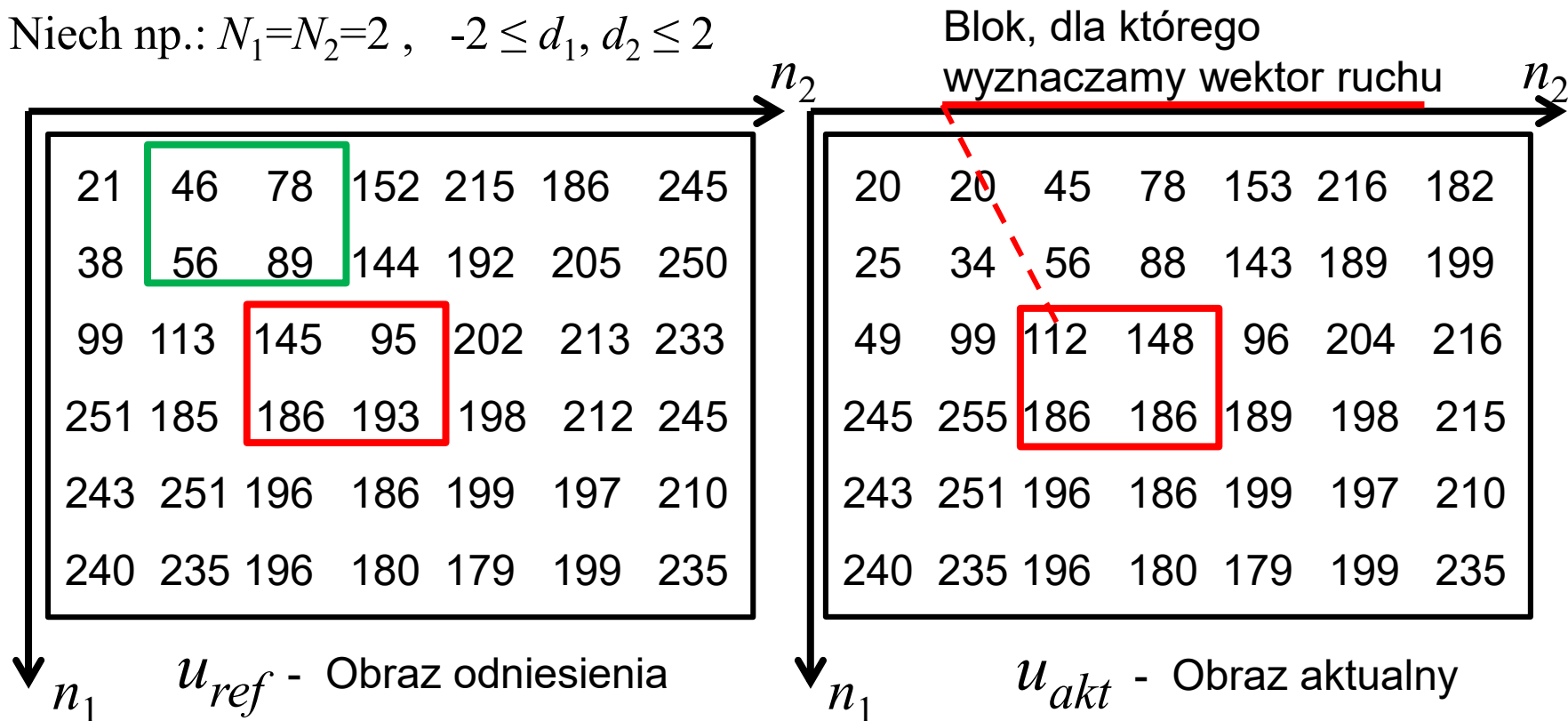
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Najpierw $d_1 = -2$ $d_2 = -2$.

$$E(2,2, n_t, -2, -2) = |112 - 21| + |148 - 46| + |186 - 38| + |186 - 56| = 471$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



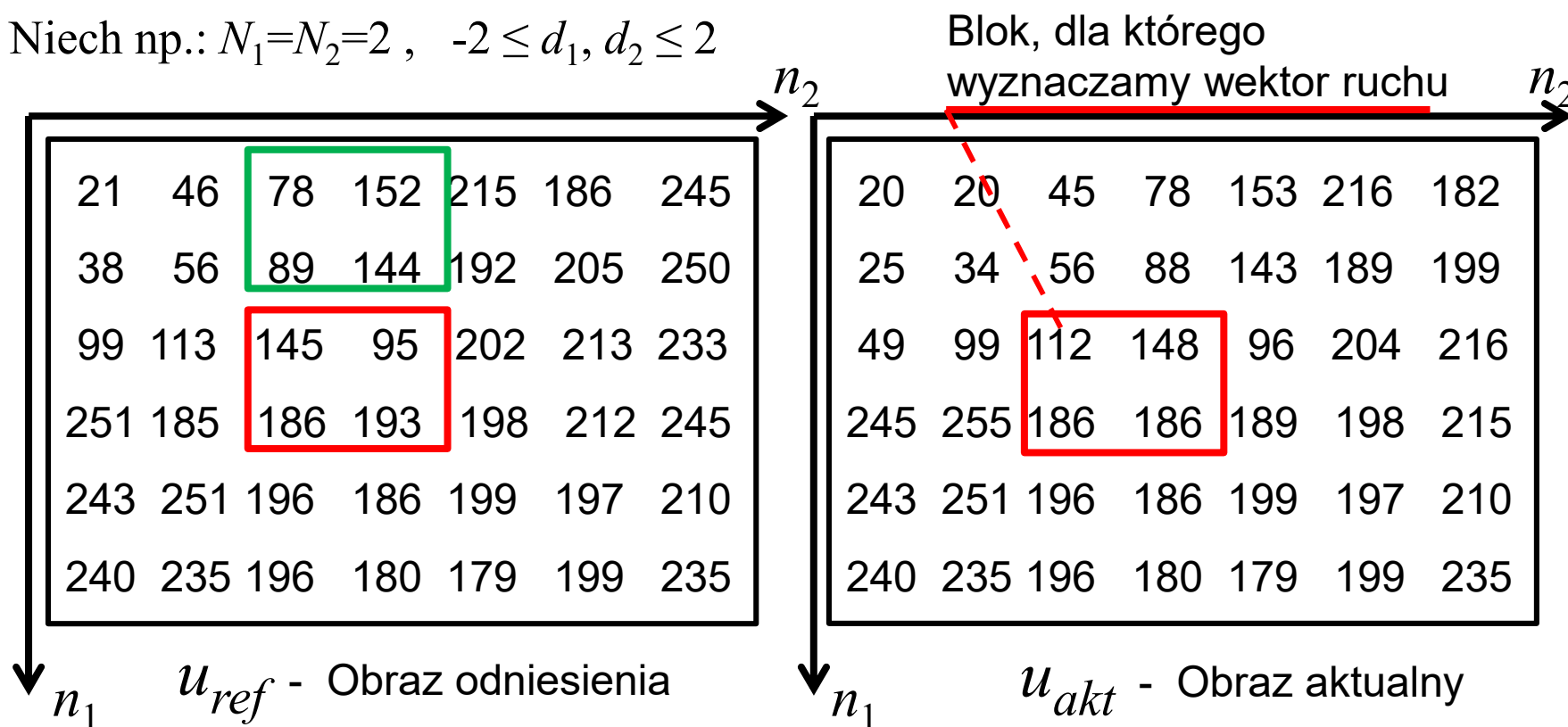
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Potem $d_1=-2$ $d_2=-1$.

$$E(2,2, n_t, -2, -1) = |112 - 46| + |148 - 78| + |186 - 56| + |186 - 89| = 66 + 70 + 130 + 97 = 363$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



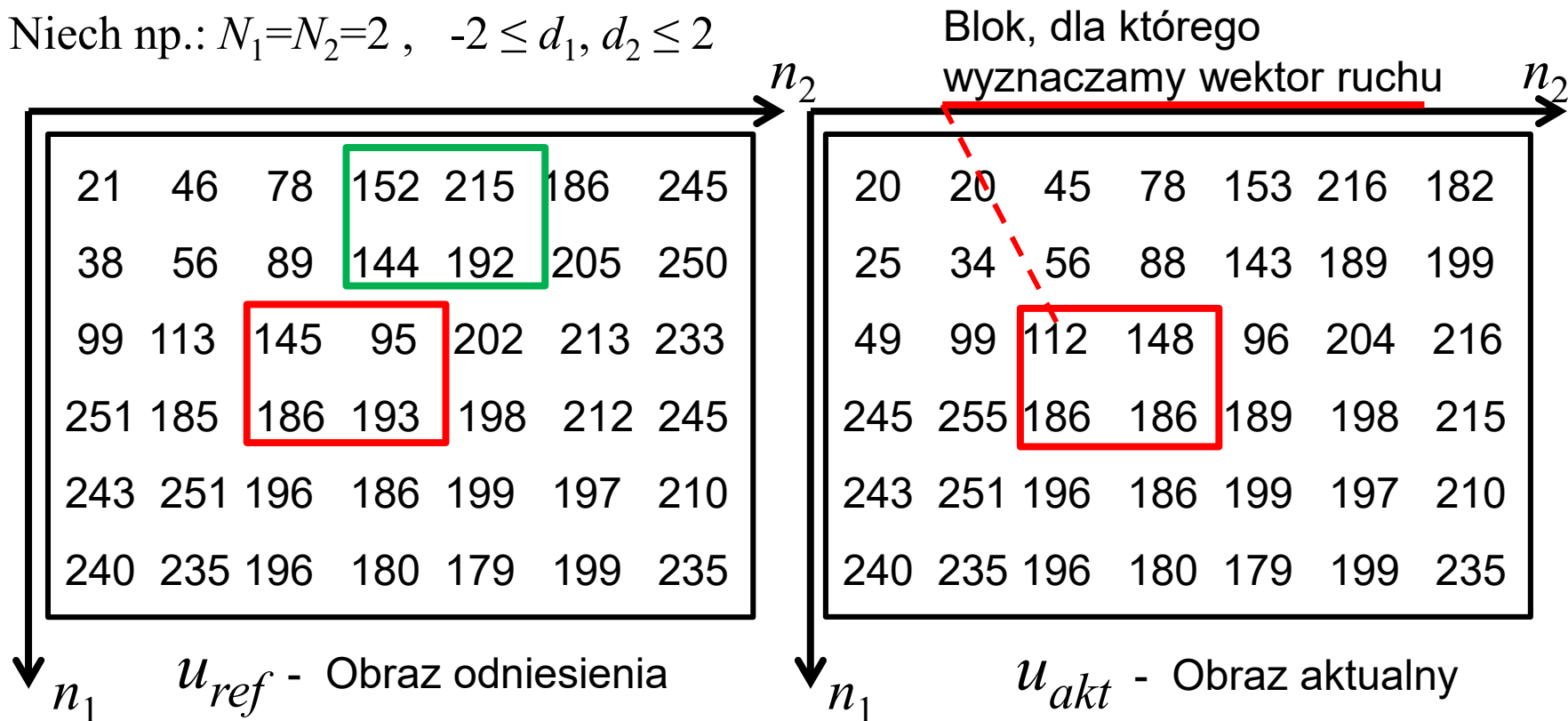
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=-2$ $d_2=0$.

$$E(2,2, n_t, -2, 0) = |112 - 78| + |148 - 152| + |186 - 89| + |186 - 144| = 34 + 4 + 97 + 42 = 177$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



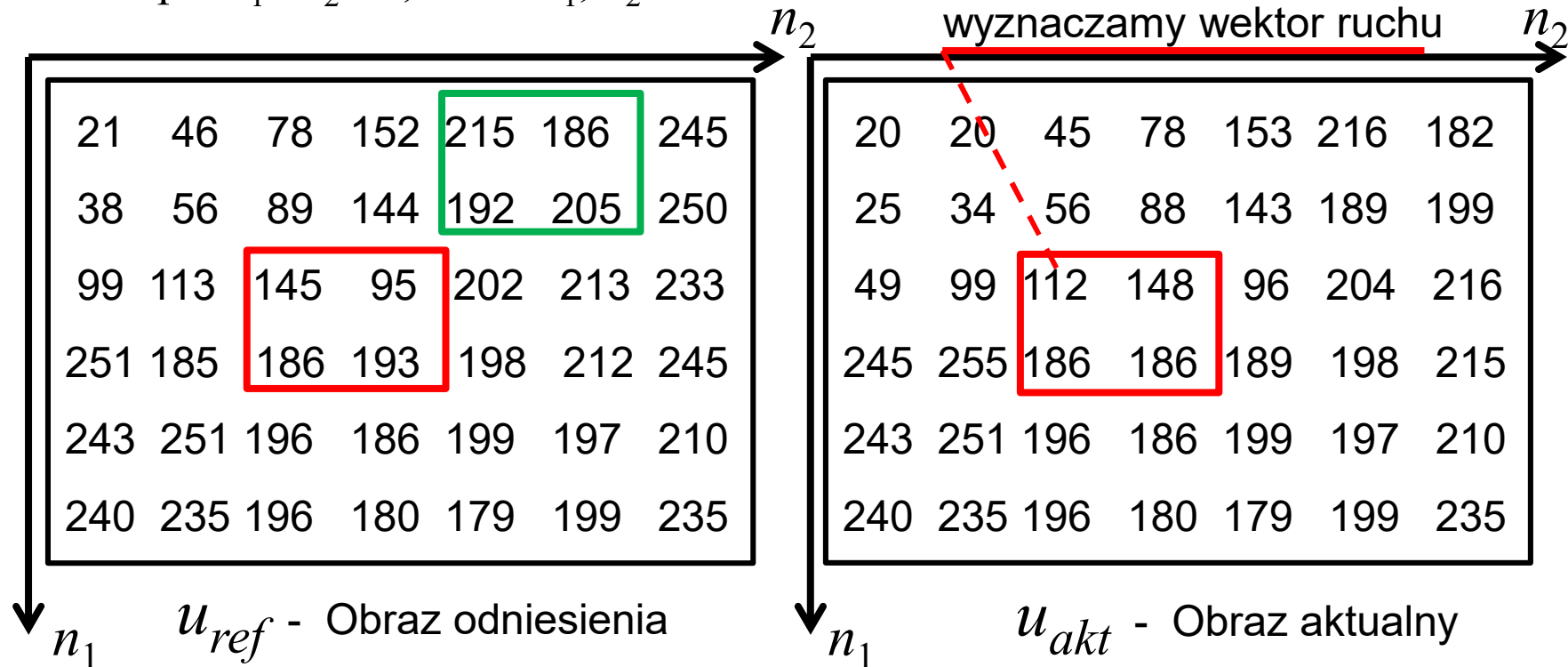
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=-2$ $d_2=1$.

$$E(2,2, n_t, -2, 1) = |112 - 152| + |148 - 215| + |186 - 144| + |186 - 192| = 40 + 67 + 42 + 6 = 155$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



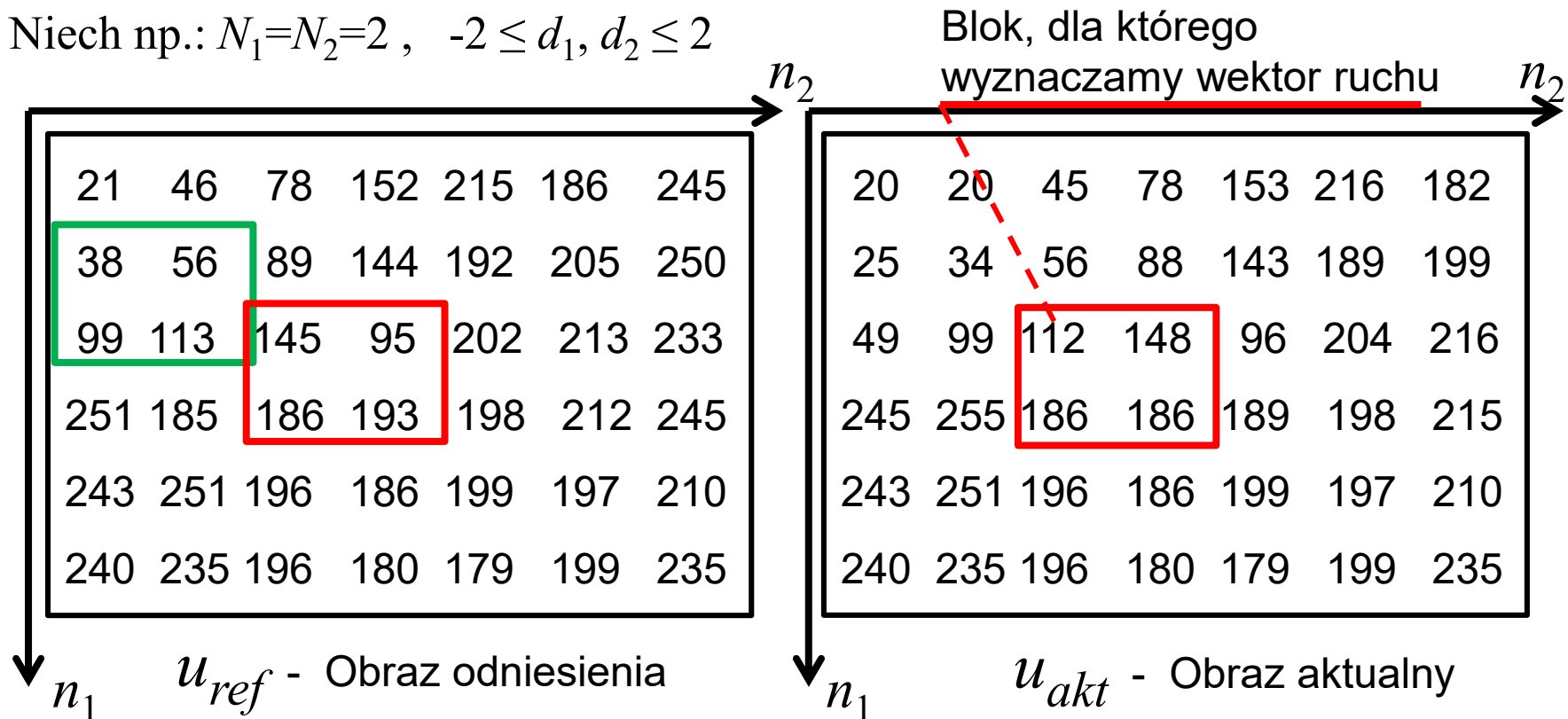
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=-2$ $d_2=2$.

$$E(2,2, n_t, -2, 2) = |112 - 215| + |148 - 186| + |186 - 192| + |186 - 205| = 103 + 38 + 6 + 19 = 166$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

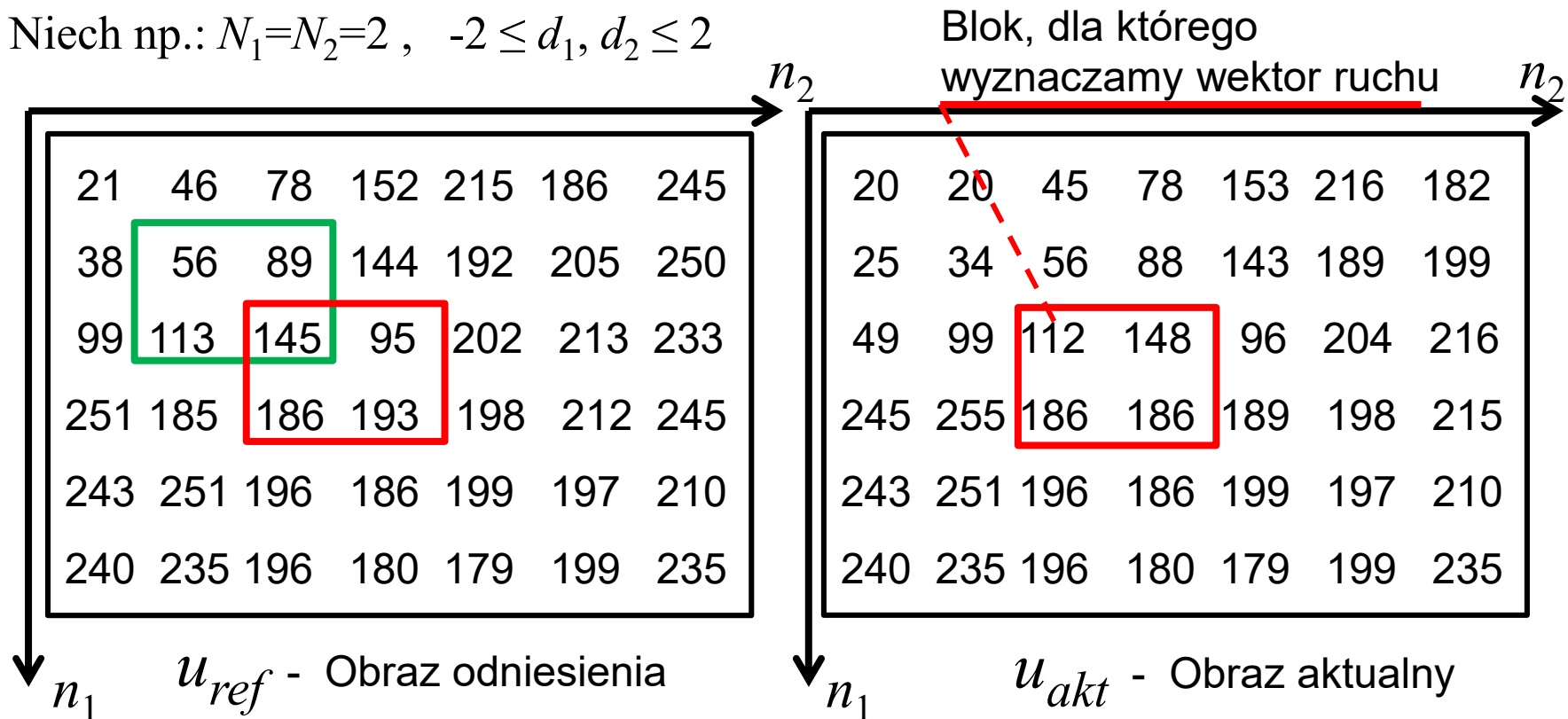


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1 = -1$ $d_2 = -2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

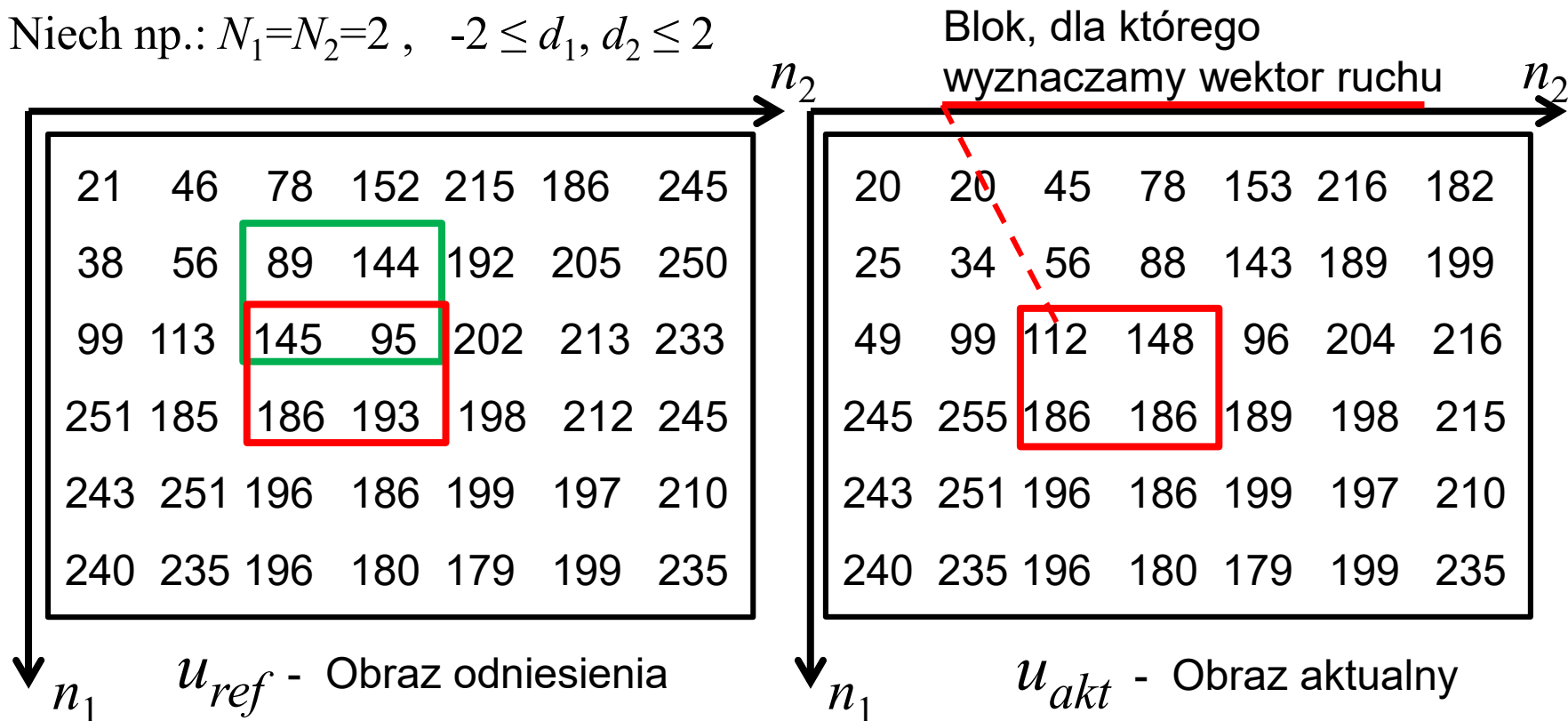


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1 = -1$ $d_2 = -1$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

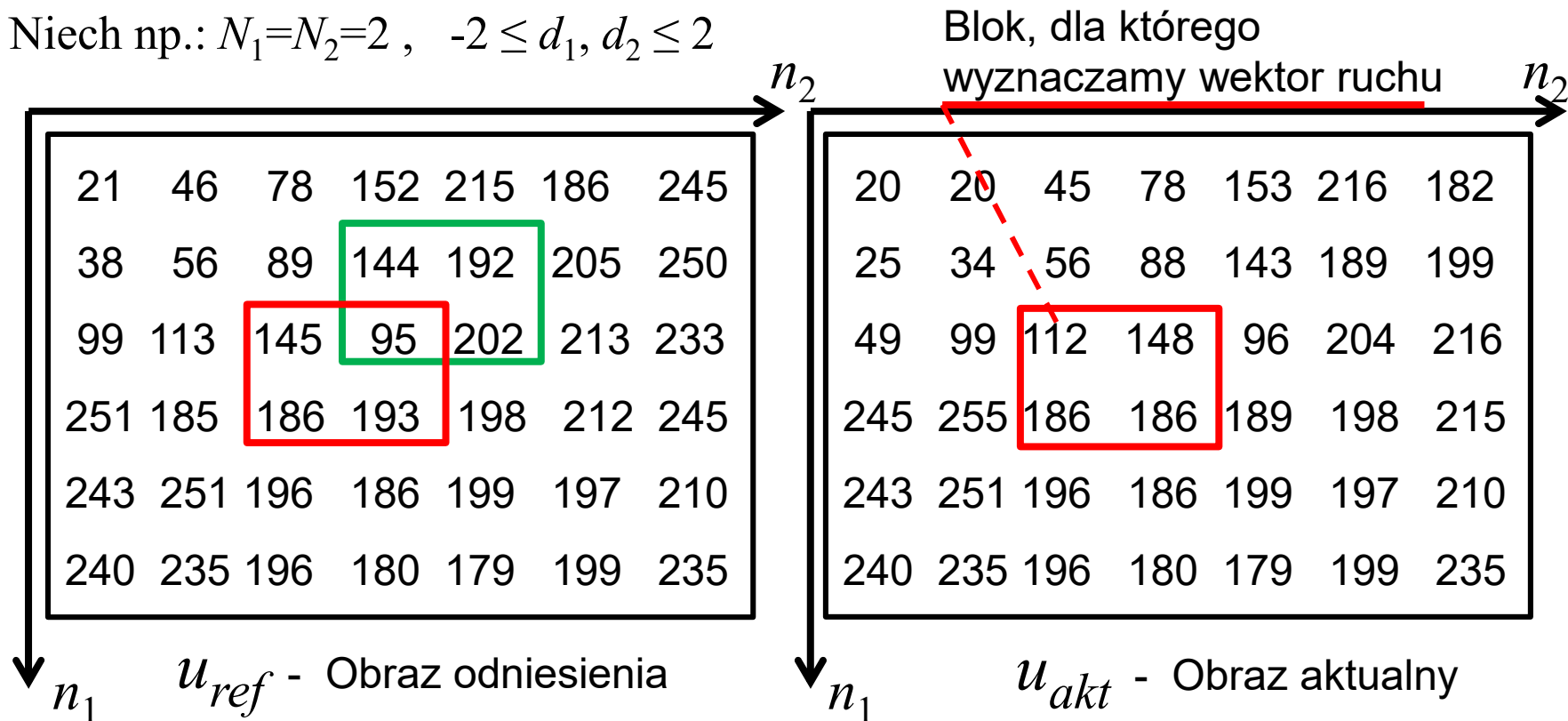


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=-1$ $d_2=0$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

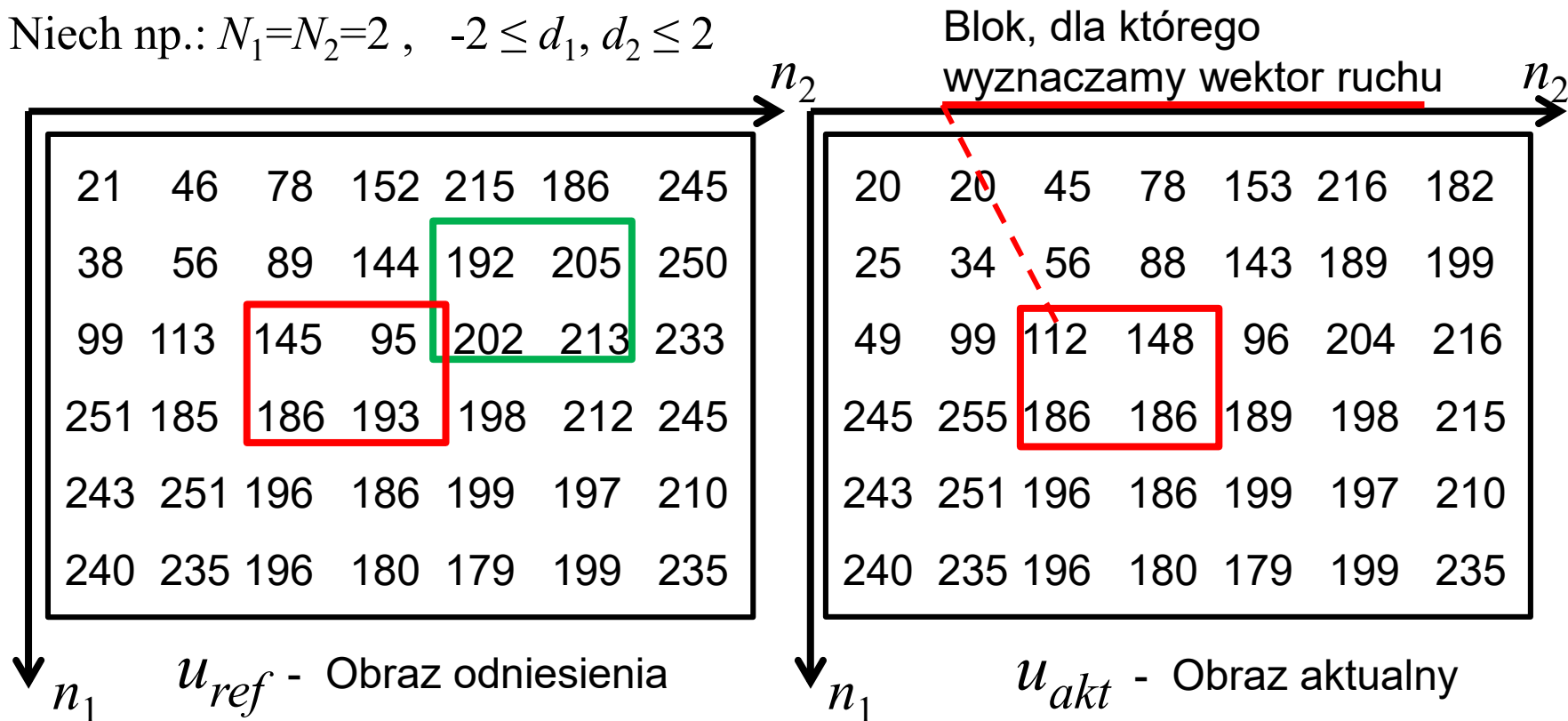


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=-1$ $d_2=1$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

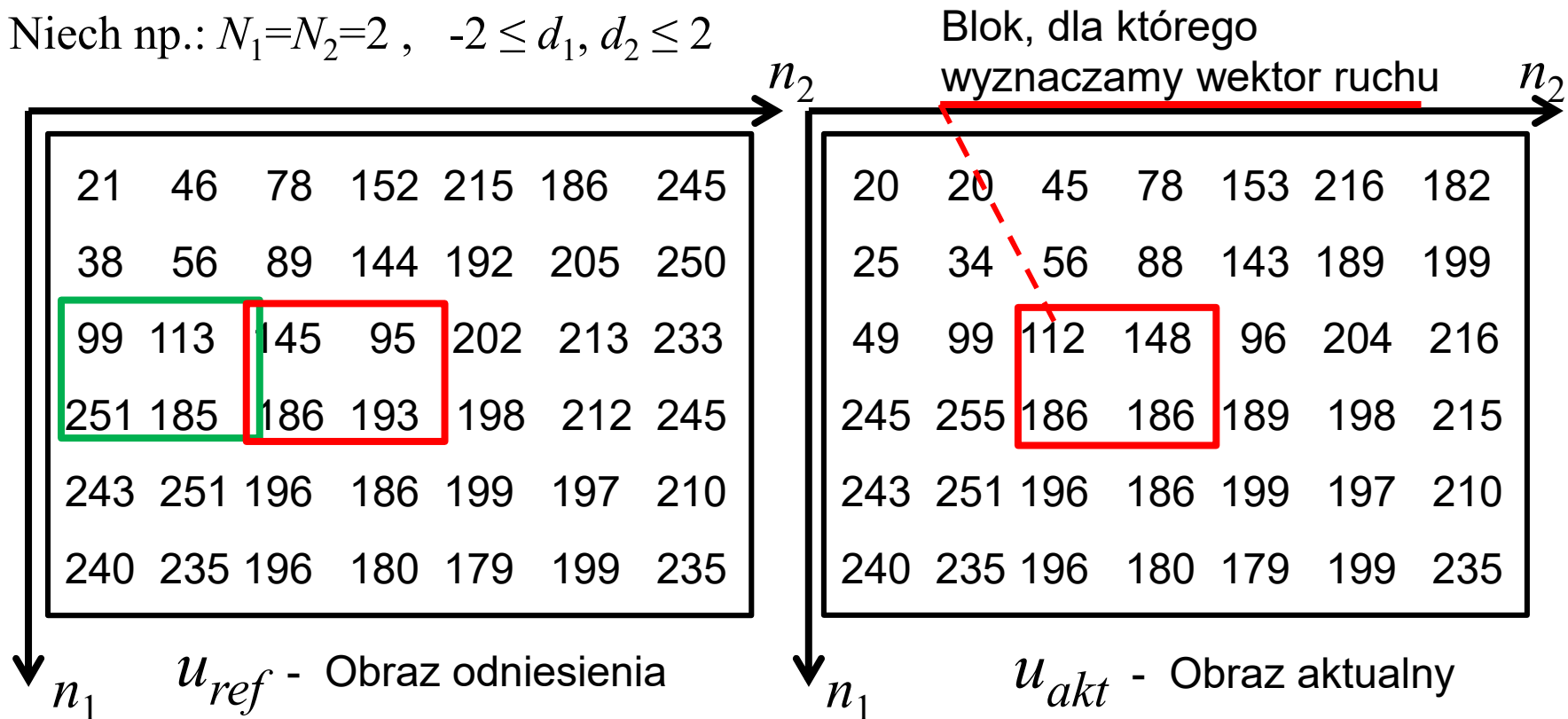


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1 = -1$ $d_2 = 2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

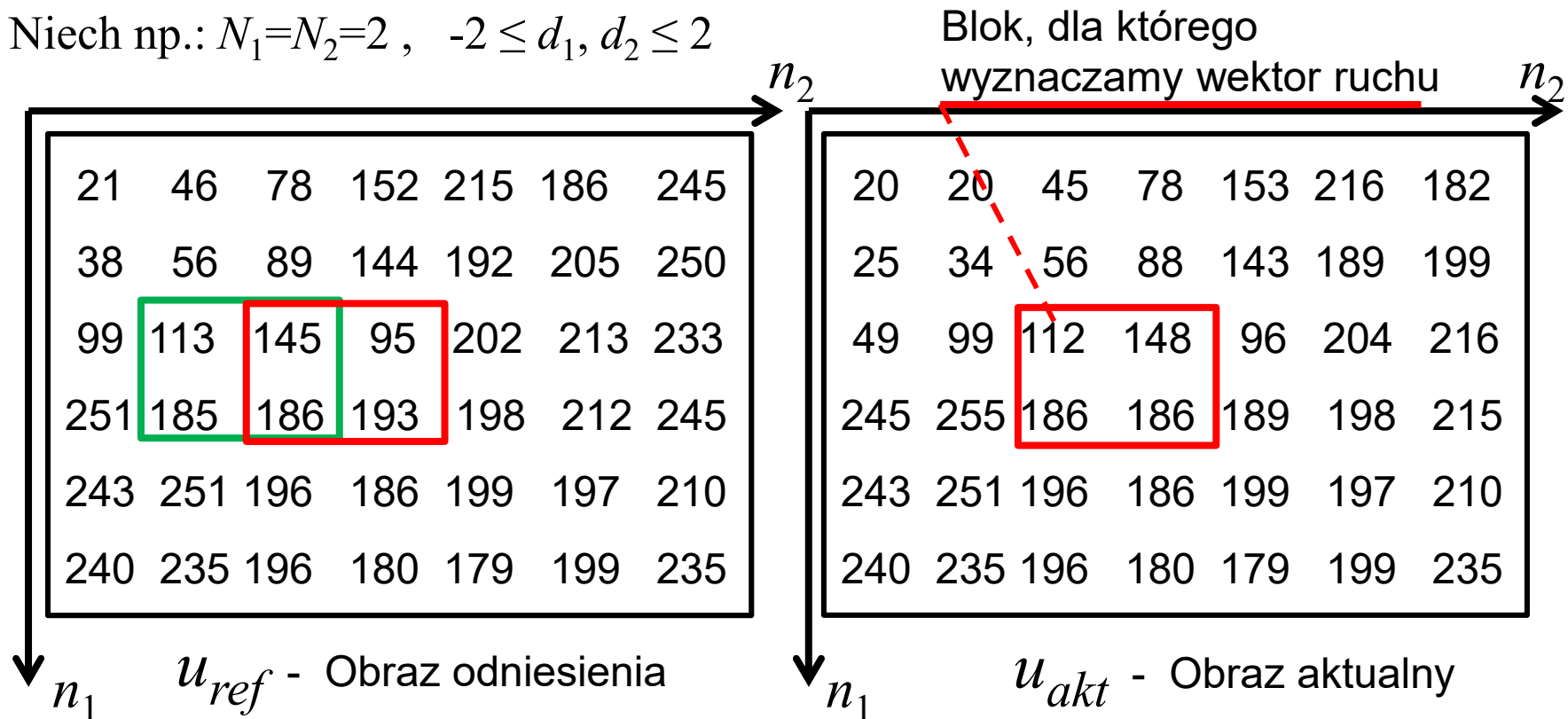


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=0$, $d_2=-2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



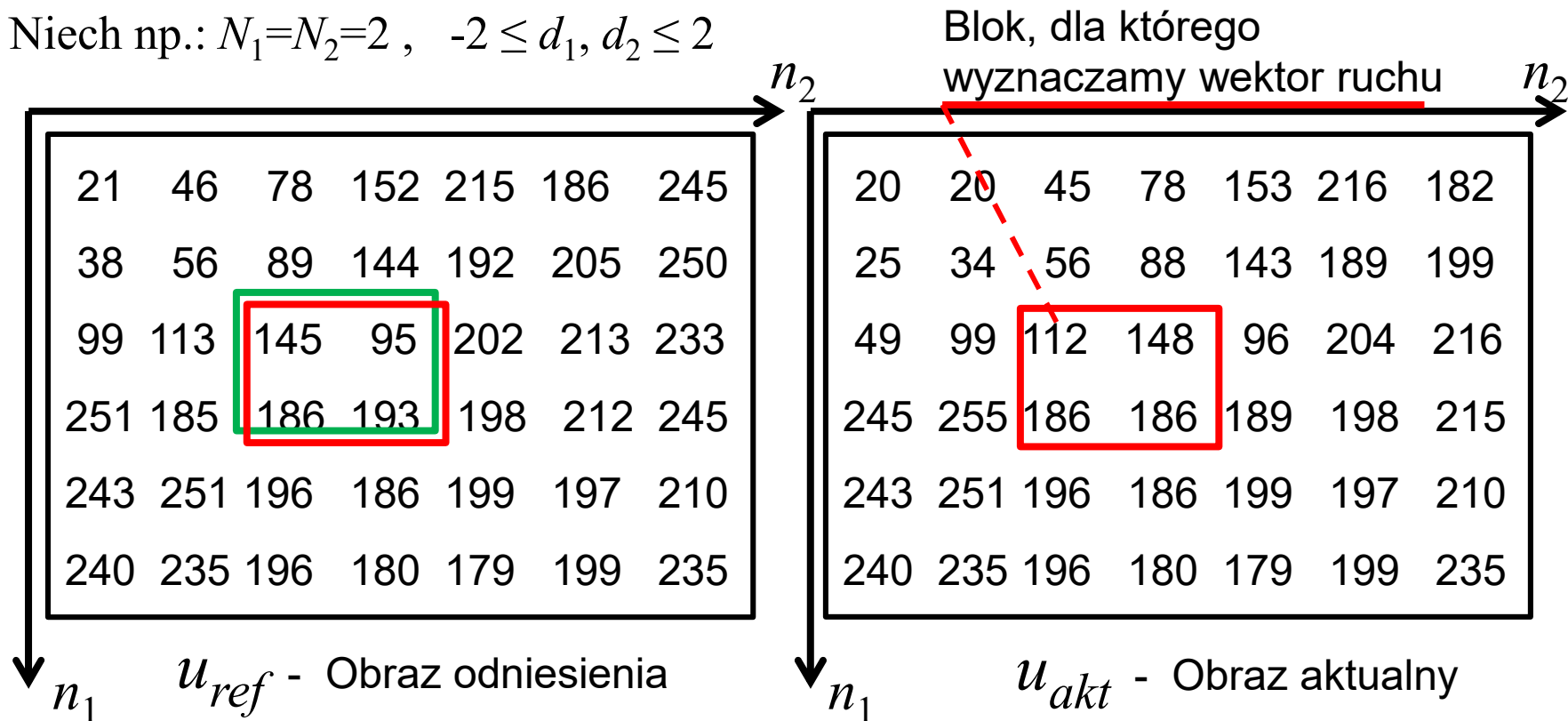
Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=0$, $d_2=-1$.

$$E(2,2, n_t, 0, -1) = |112 - 113| + |148 - 145| + |186 - 185| + |186 - 186| = 1 + 3 + 1 + 0 = 5$$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

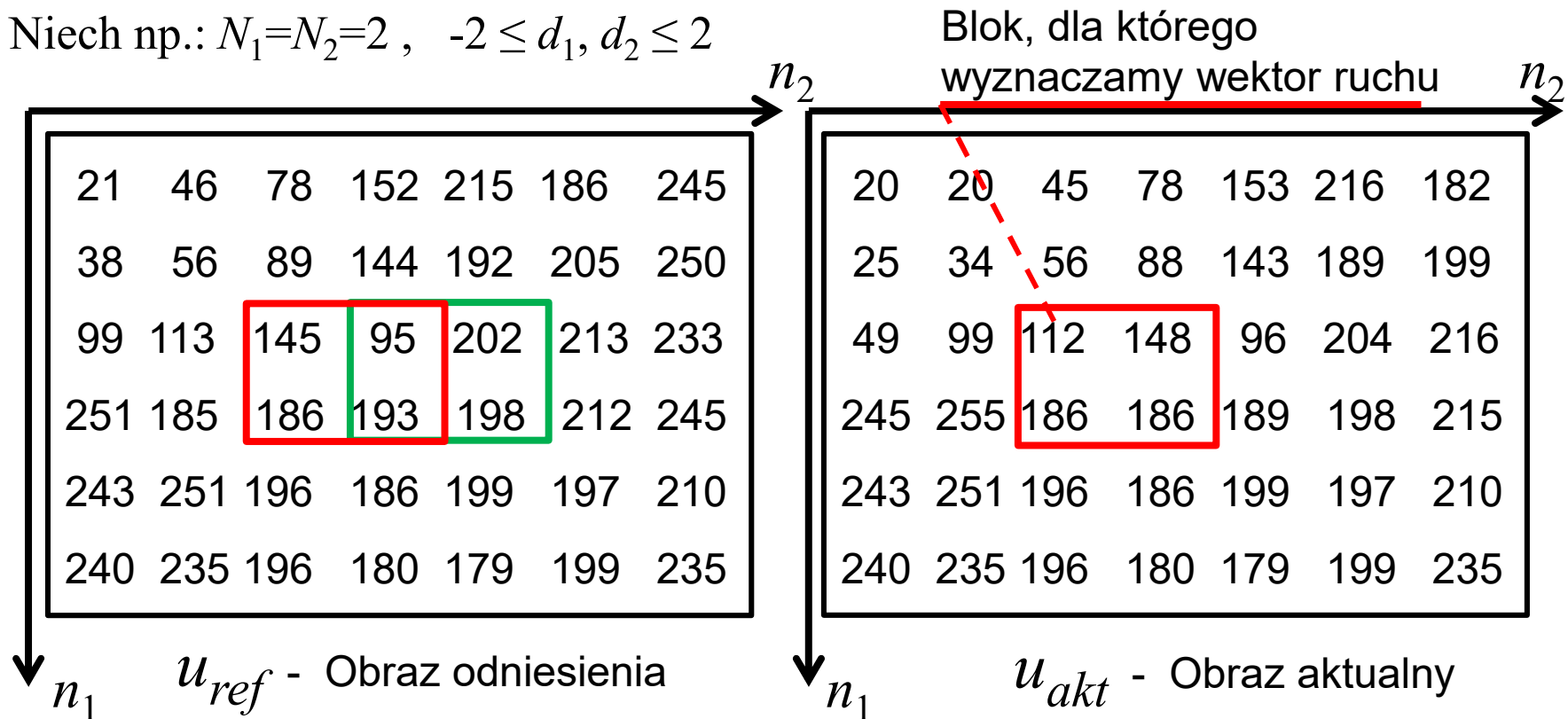


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=0$, $d_2=0$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

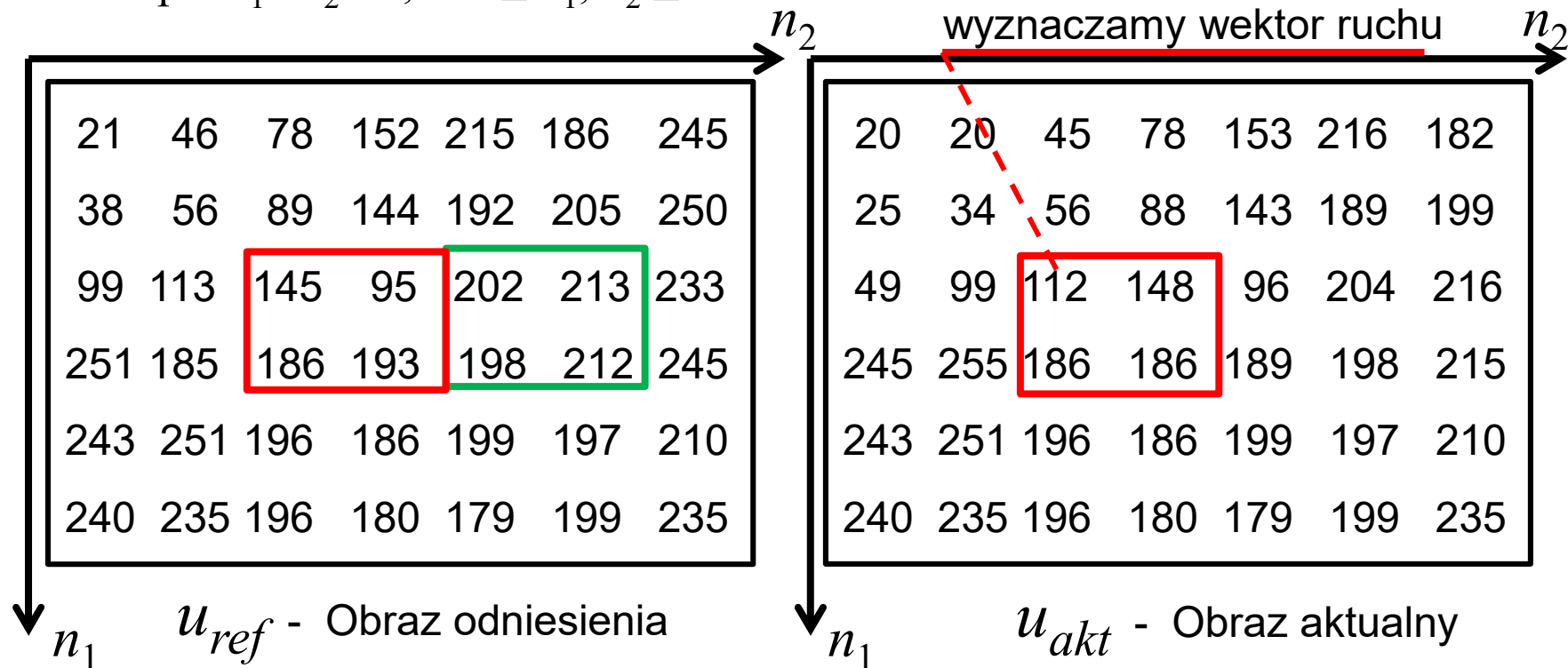


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=0$, $d_2=1$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

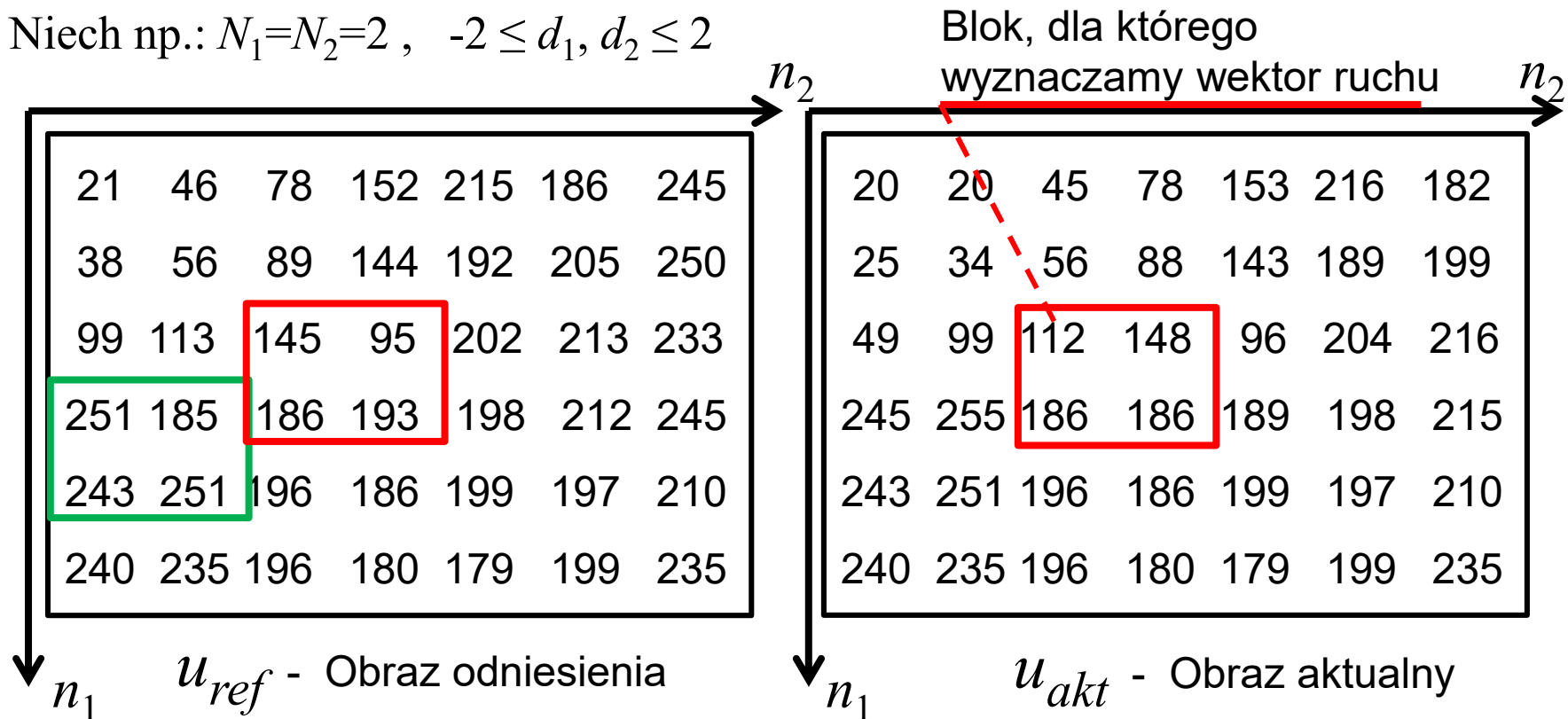


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=0$, $d_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

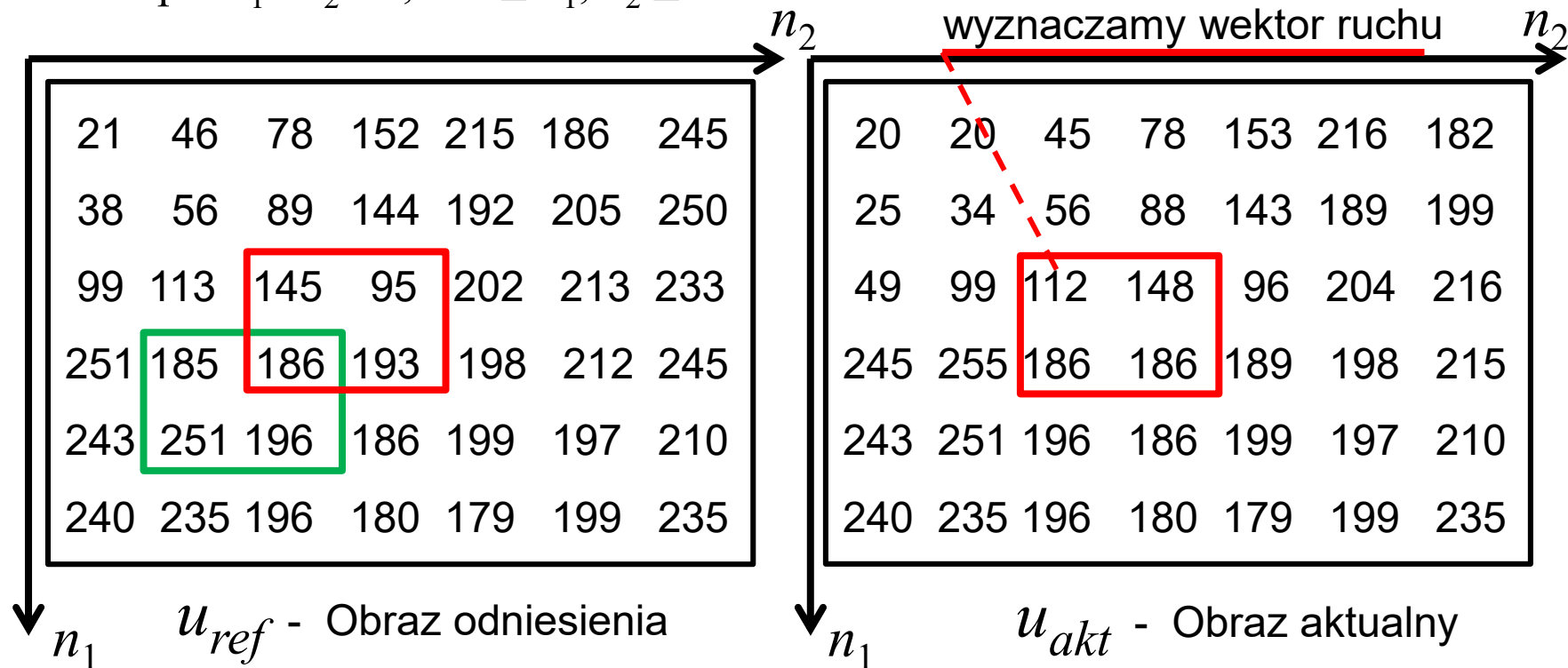


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$. Teraz $d_1=1$, $d_2=-2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

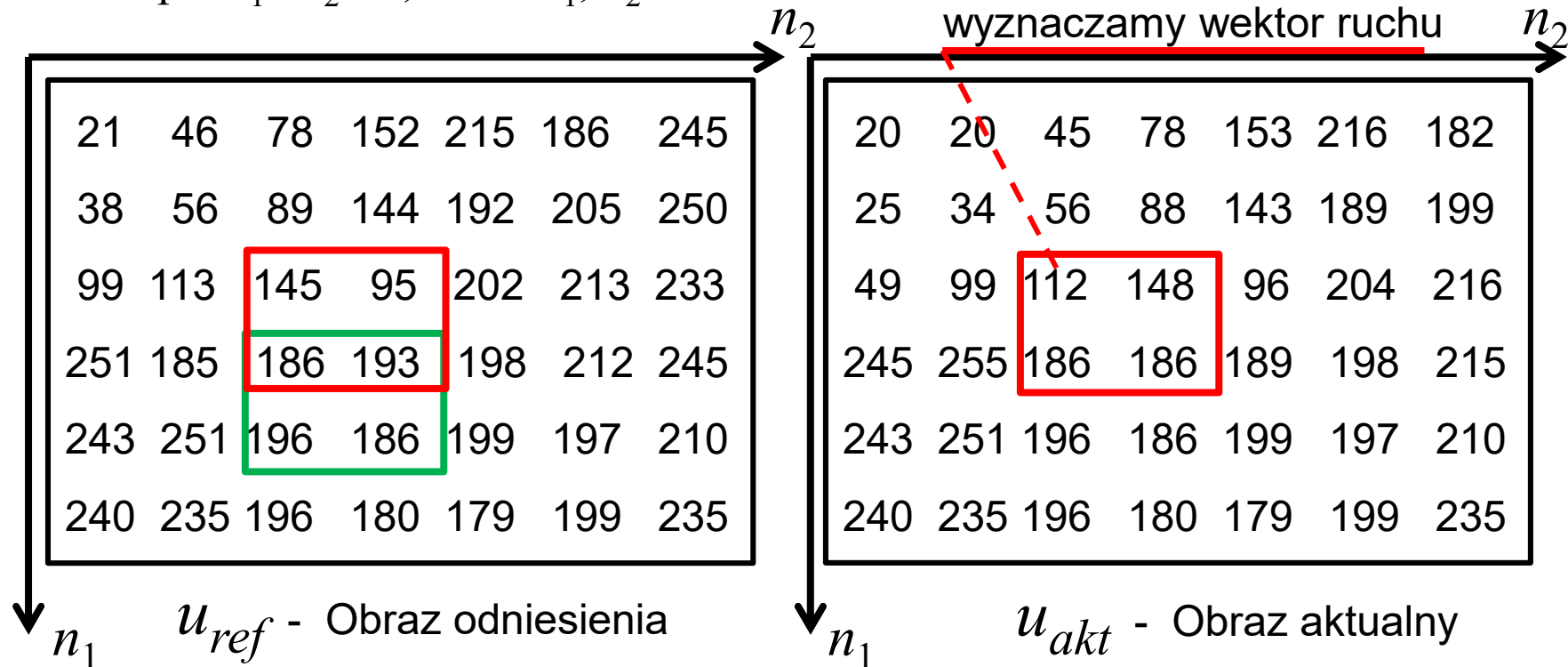


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

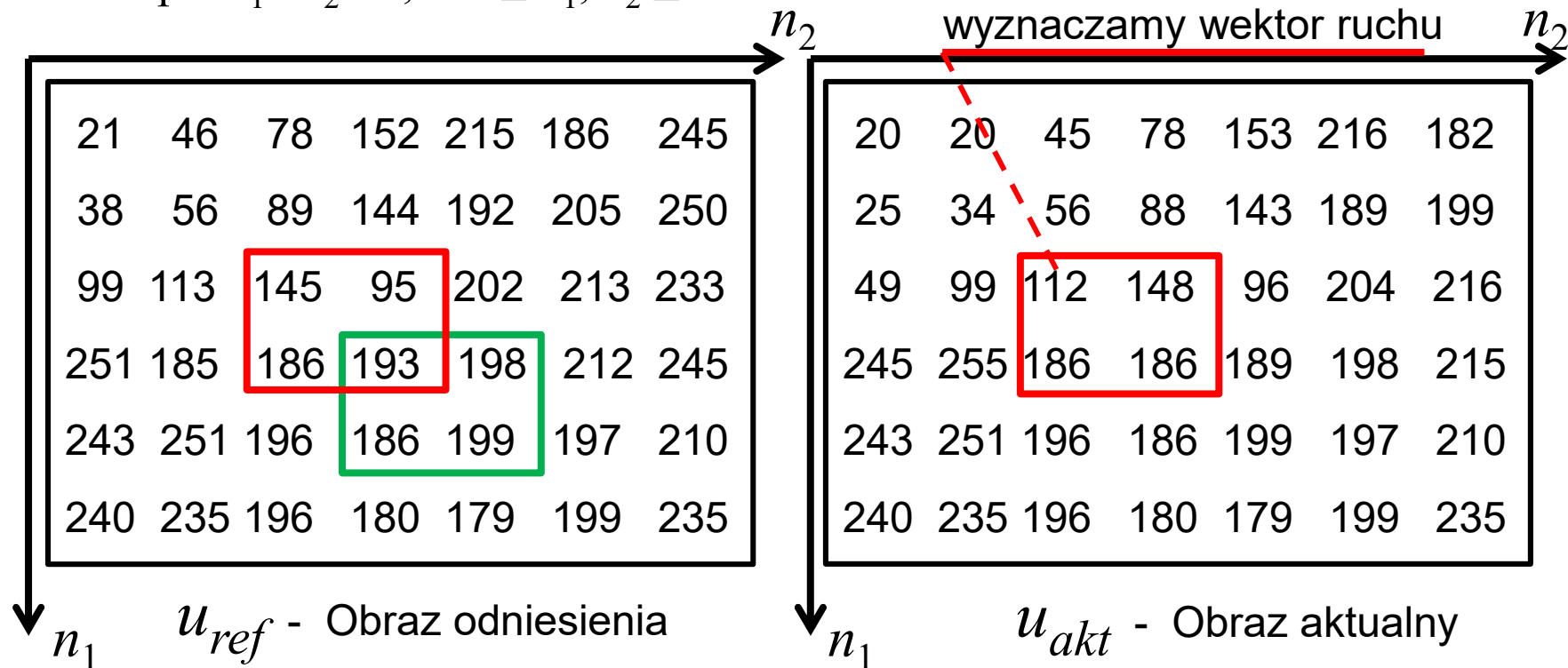


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

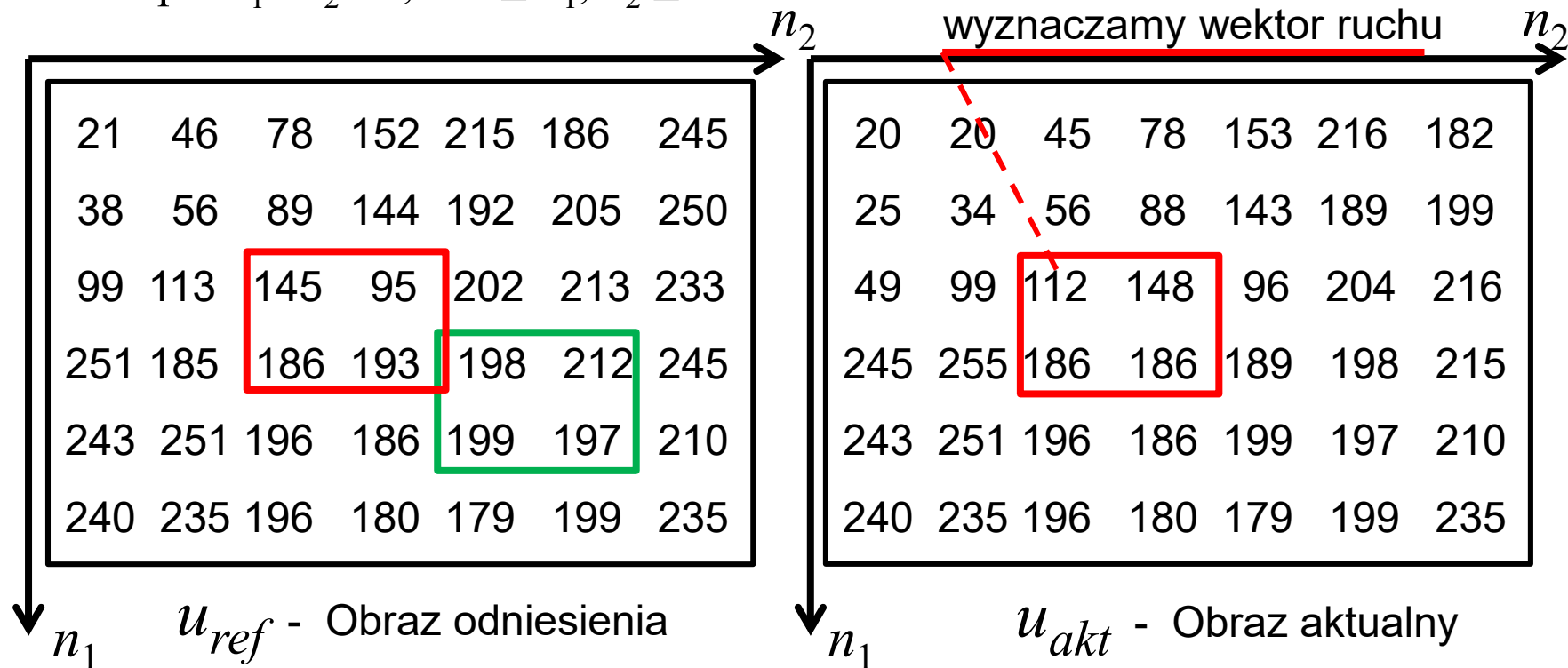


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

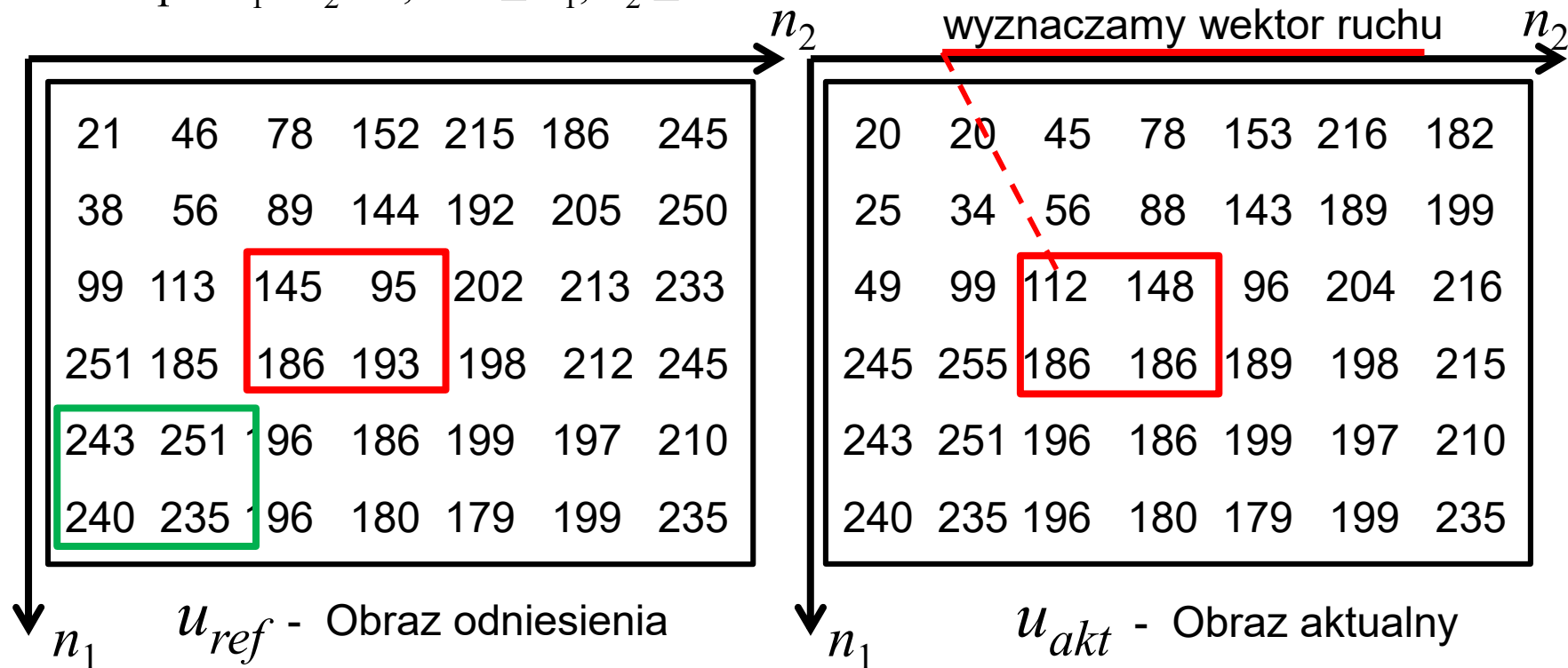


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

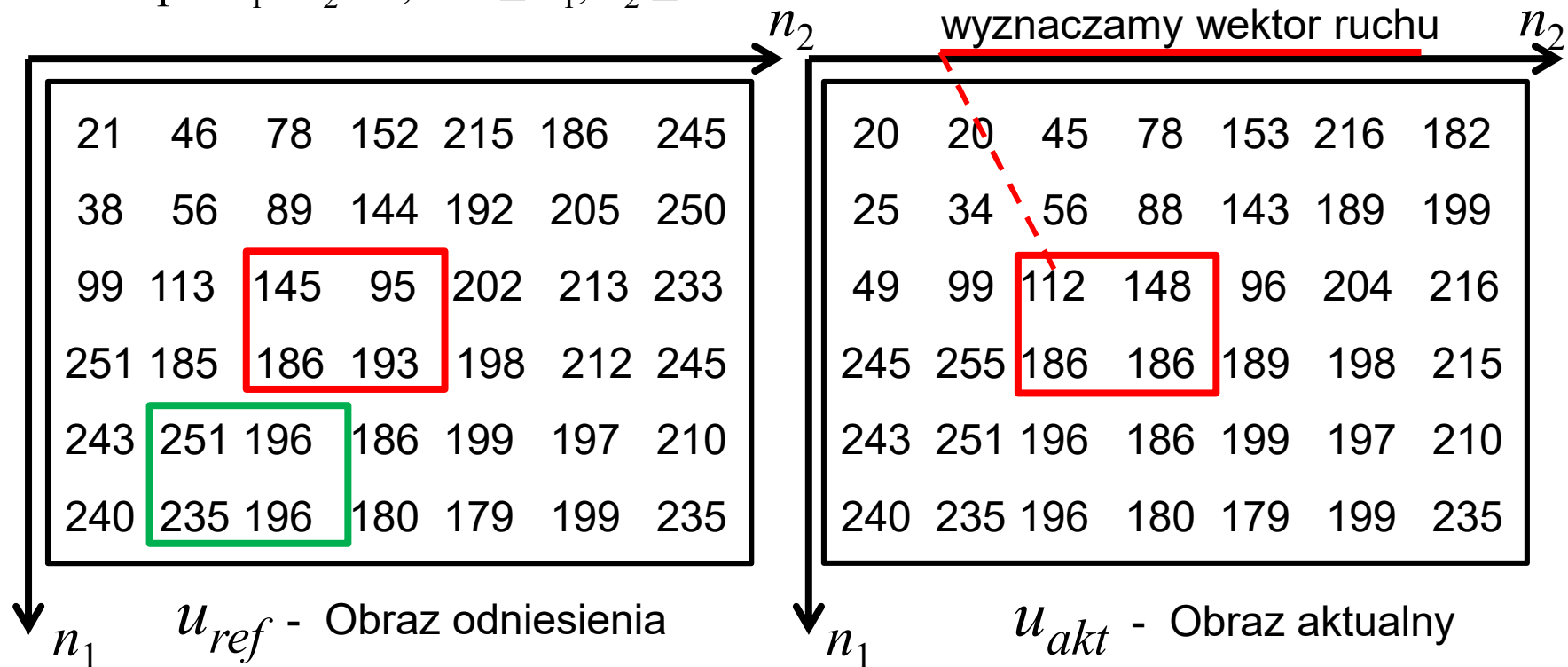


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

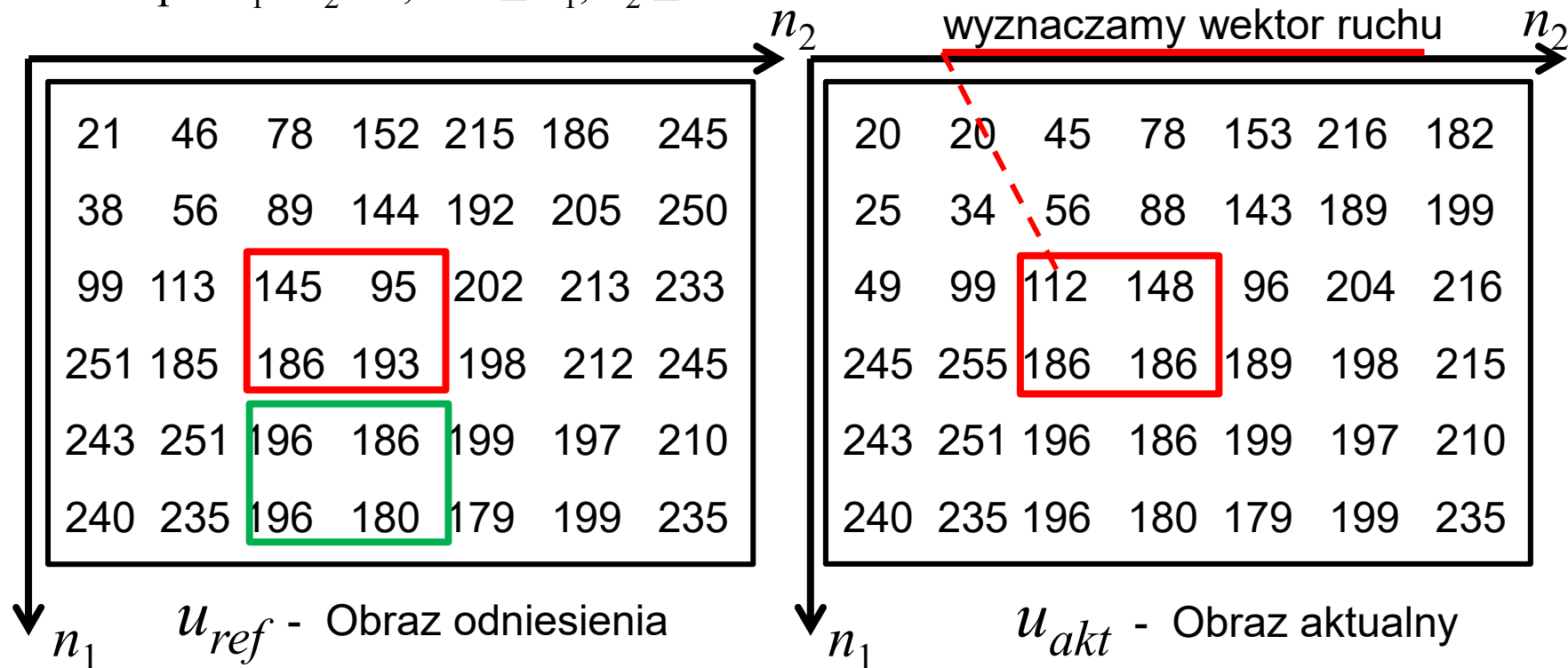


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

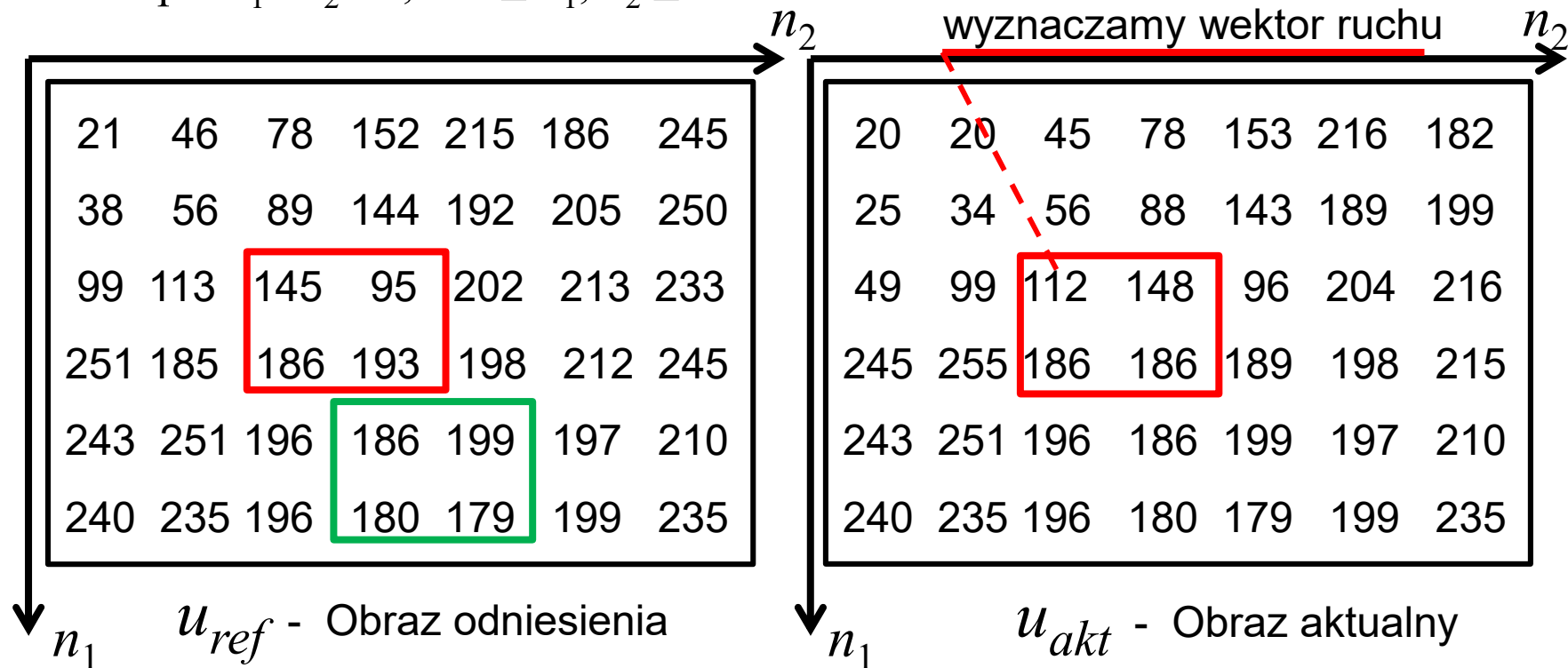


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$

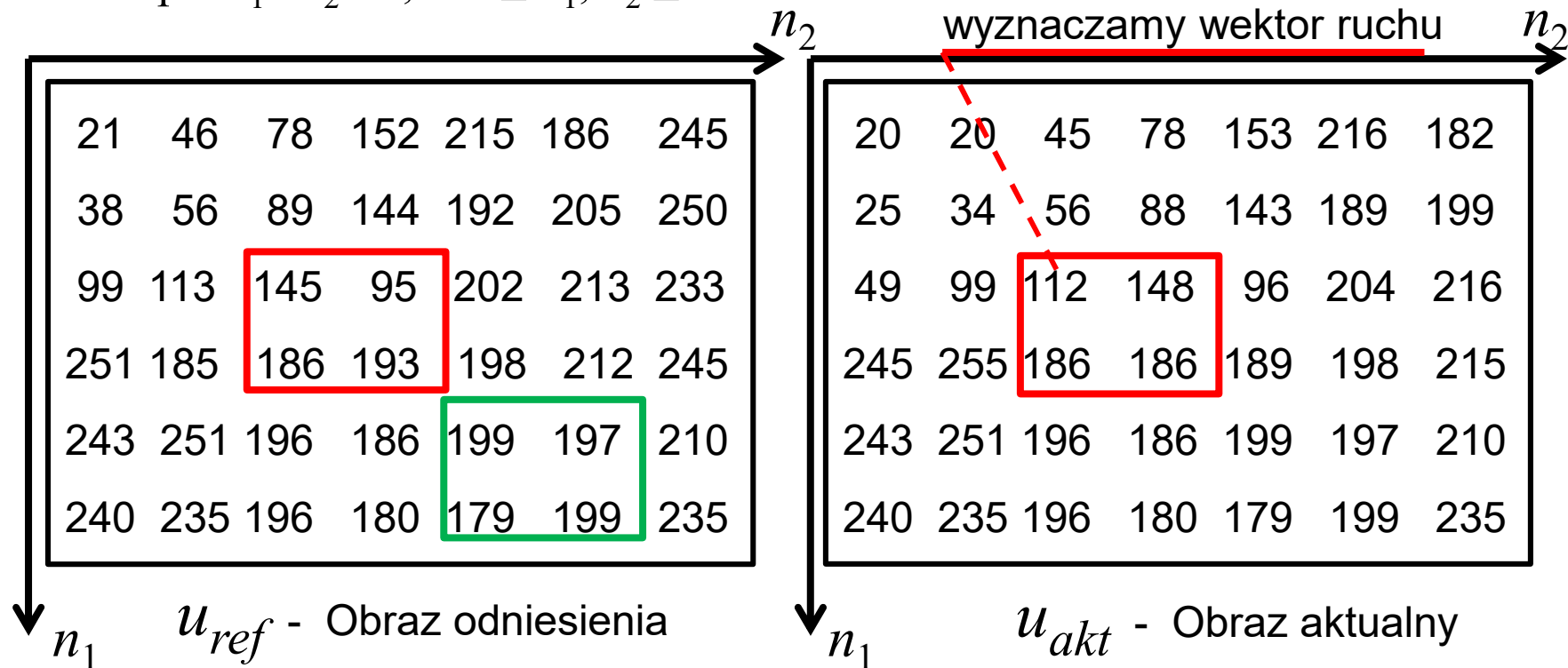


Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



Wektor ruchu wyznaczamy dla bloku $n_1=n_2=2$.

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Założyliśmy: $-2 \leq d_1, d_2 \leq 2$

Porównujemy

$E(-2,-2)= 471, E(-2,-1)= 363, E(-2,0)= 177, E(-2,1)= 155, E(-2,2)= 19,$

$E(-1,-2)= ..., ...$

$E(0,-2)= ..., E(0,-1)= 5, E(0,0)= ...,$

$...,$

$E(2,-2)= ...$

$E(2,2)= ...,$

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \left| u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t) \right|,$$

Założyliśmy: $-2 \leq d_1, d_2 \leq 2$

Porównujemy

$E(-2,-2)= 471, E(-2,-1)= 363, E(-2,0)= 177, E(-2,1)= 155, E(-2,2)= 19,$

$E(-1,-2)= \dots, \dots$

$E(0,-2)= \dots, E(0,-1)= 5, E(0,0)= \dots, \dots$

$\dots,$

$E(2,-2)= \dots$

$E(2,2)= \dots,$

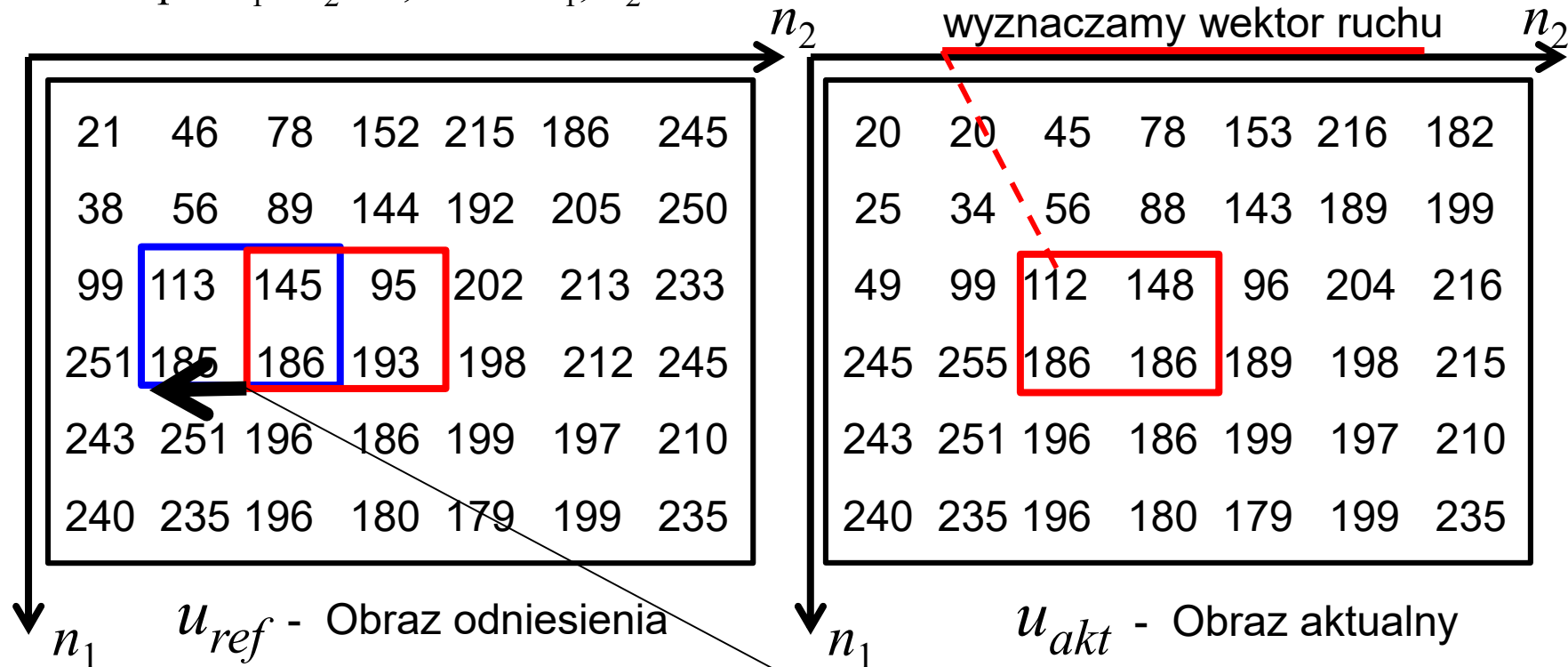


Wektor ruchu = (0, -1)

Estymacja ruchu - przykład

$$E(n_1, n_2, n_t; d_1, d_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |u_{akt}(n_1 + k_1, n_2 + k_2, n_t) - u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)|,$$

Niech np.: $N_1=N_2=2$, $-2 \leq d_1, d_2 \leq 2$



Najbardziej podobny blok
uzyskany dla $SAD=5$: $d_1=0$, $d_2=-1$.

WEKTOR RUCHU

Estymacja ruchu (c.d.)

Metoda pasowania bloków

Poszukiwana jest para liczb (d_1, d_2) minimalizująca wskaźnik

$$E(n_1, n_2, n_t; d_1, d_2) = \\ = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} F[u_{akt}(n_1 + k_1, n_2 + k_2, n_t), u_{ref}(n_1 + k_1 + d_1, n_2 + k_2 + d_2, n_t - M_t)],$$

Przykład: $N_1=N_2=16$ (makroblok MB)

16 x 16 = 256 porównań punktów dwóch bloków 16 x 16

x 65 x 65 = 1081600 obszar przeszukiwania $-32 < d < 32$

x 36 x 45 = 1752192000 liczba MB w obrazie SDTV

x 25 = 4,38 · 10¹⁰ 25 obrazów na sekundę

Wektory ruchu (przykład)



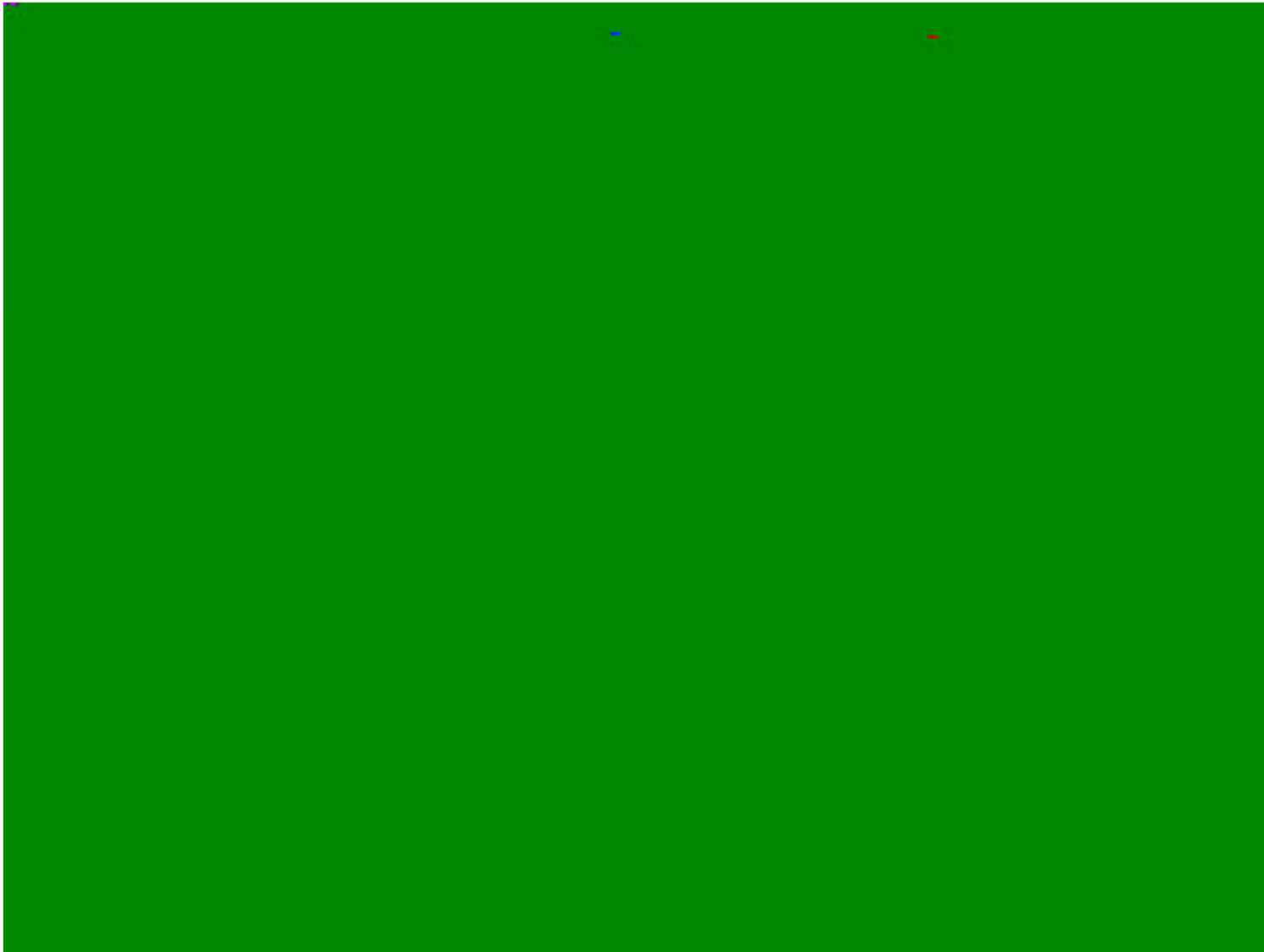
Wektory ruchu (przykład)



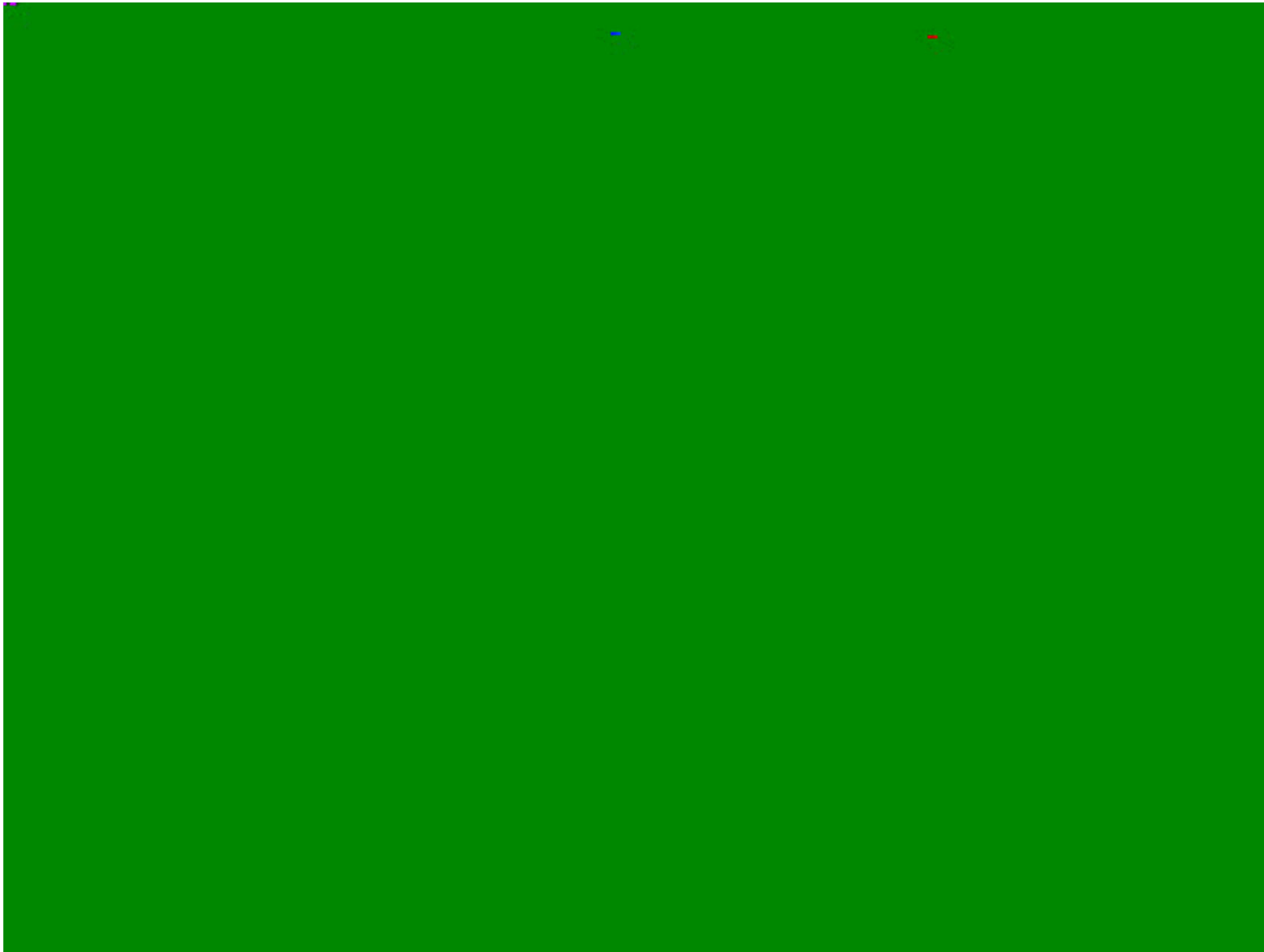
Wektory ruchu (przykład)



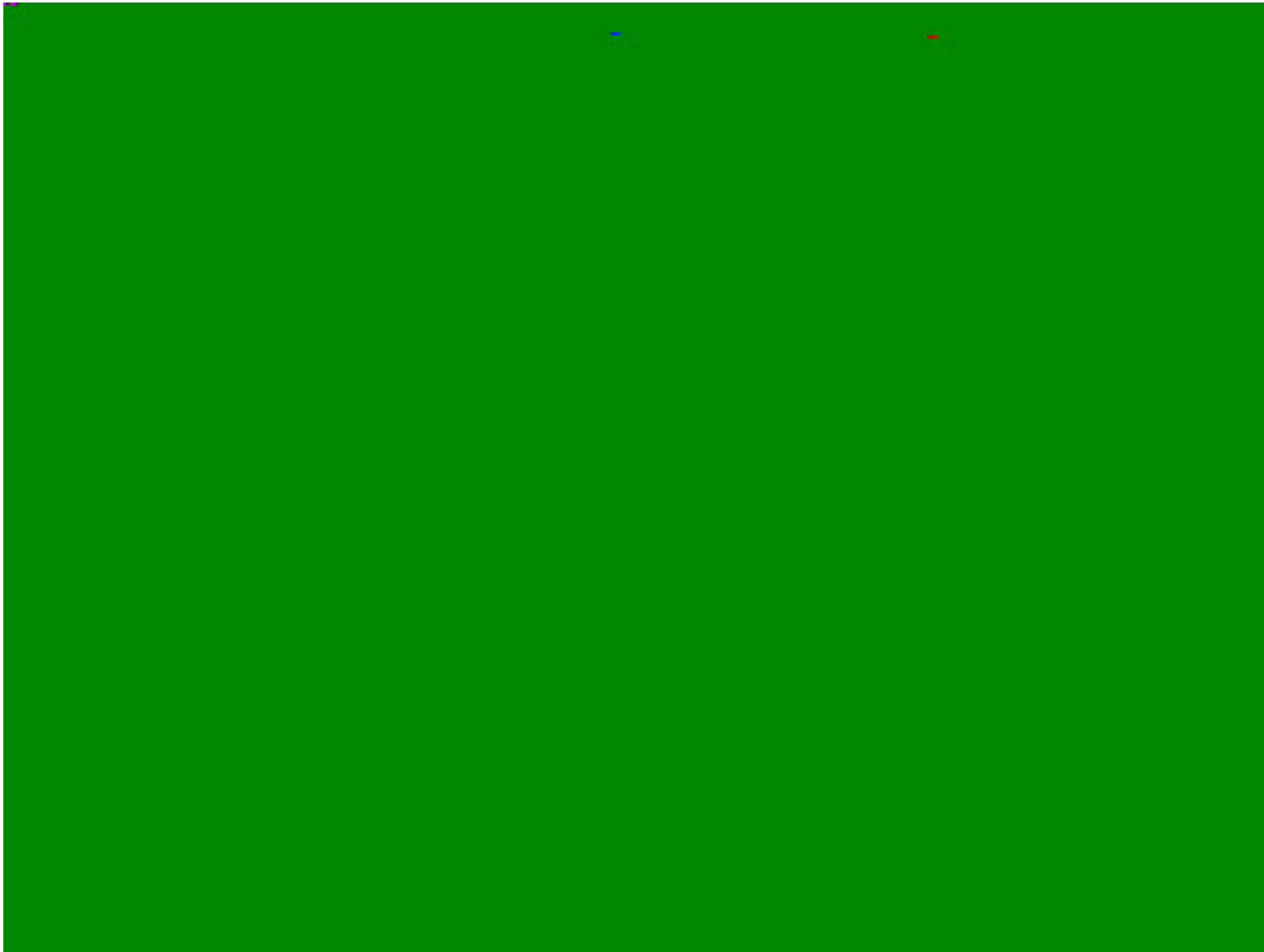
Wektory ruchu (przykład)



Wektory ruchu (przykład)



Wektory ruchu (przykład)



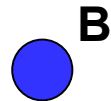
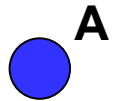
Dokładność estymacji ruchu

- W nowszych technikach kompresji stosuje się estymację ruchu z dokładnością do $1/4$ okresu próbkowania .

Znaczenie dokładności „ułamkowej” w estymacji ruchu - przykład



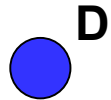
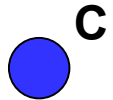
Dokładność do ułamków okresu próbkowania („*subpixel accuracy*”)



$$a = (A + B + 1)/2$$



$$b = (A + C + 1)/2$$



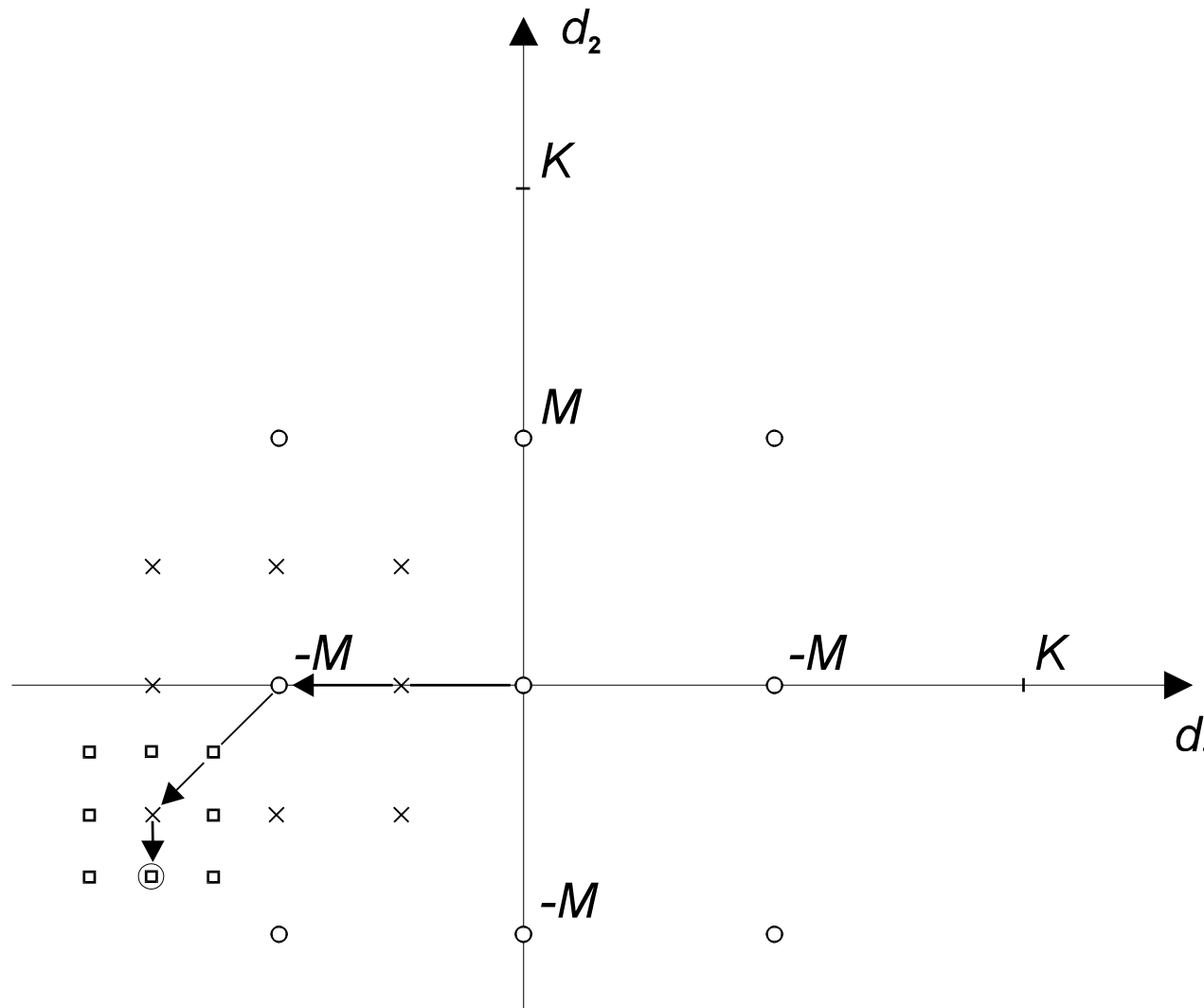
$$c = (A + B + C + D + 2)/4$$

Estymacja ruchu

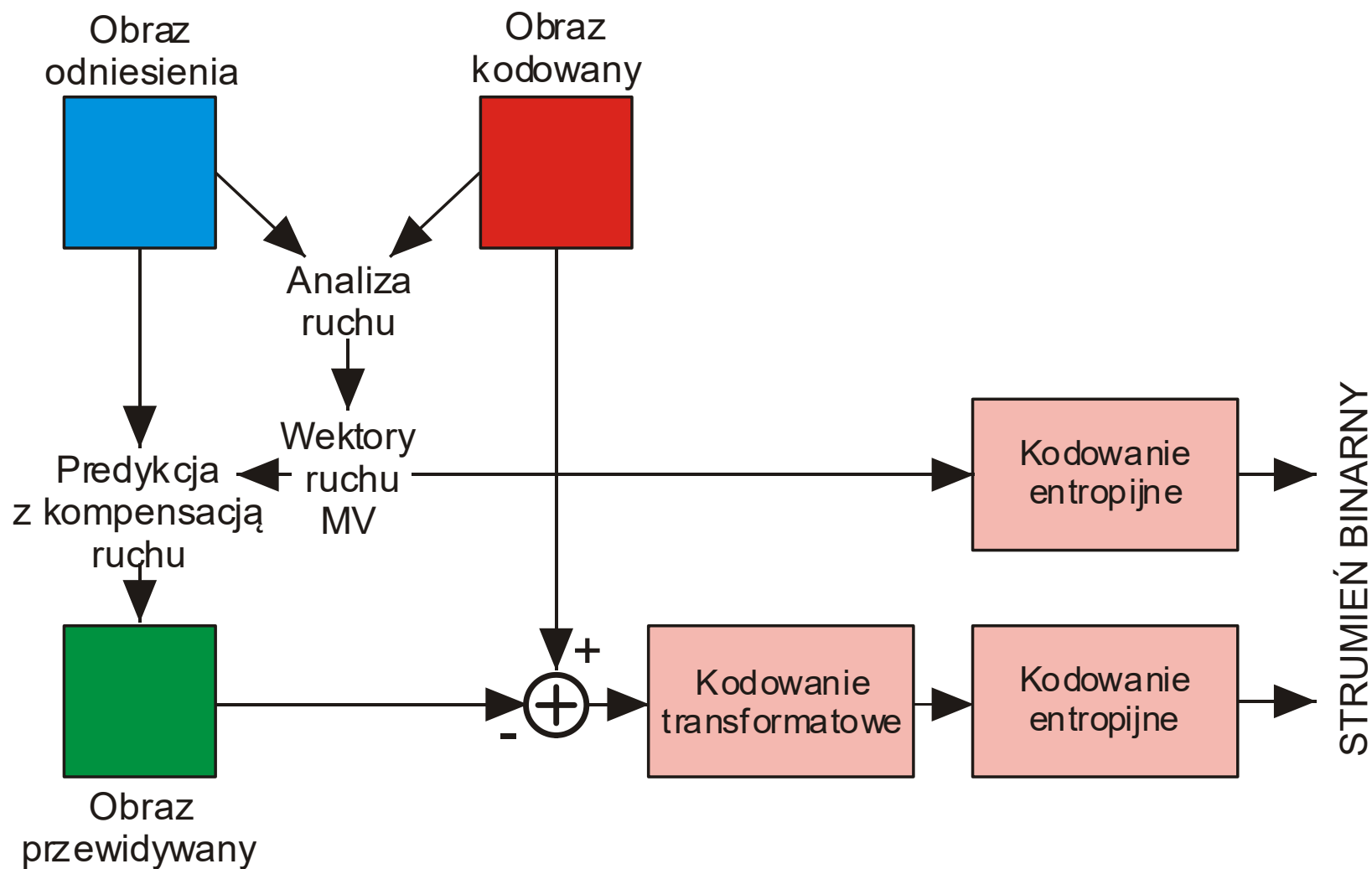
- Najbardziej pracochłonna operacja w koderze.
- Około 40-70% mocy obliczeniowej koderu.
- Dąży się do redukcji mocy obliczeniowej przez zastosowanie **szybkich metod estymacji ruchu** zamiast **pełnego przeszukiwania**.

Estymacja ruchu (c.d.) – przeszukiwanie logarytmiczne

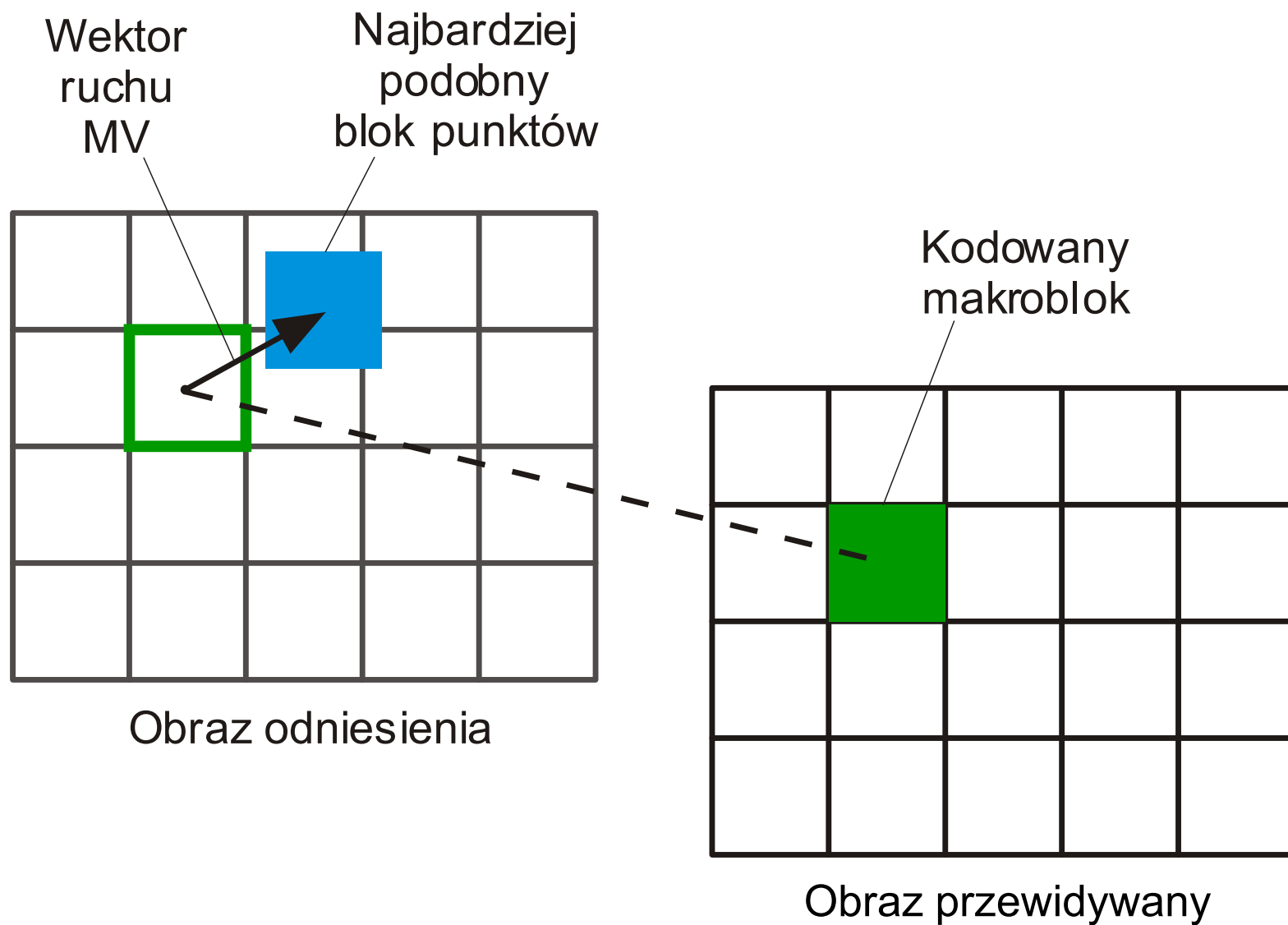
Przykład szybkiej techniki estymacji ruchu



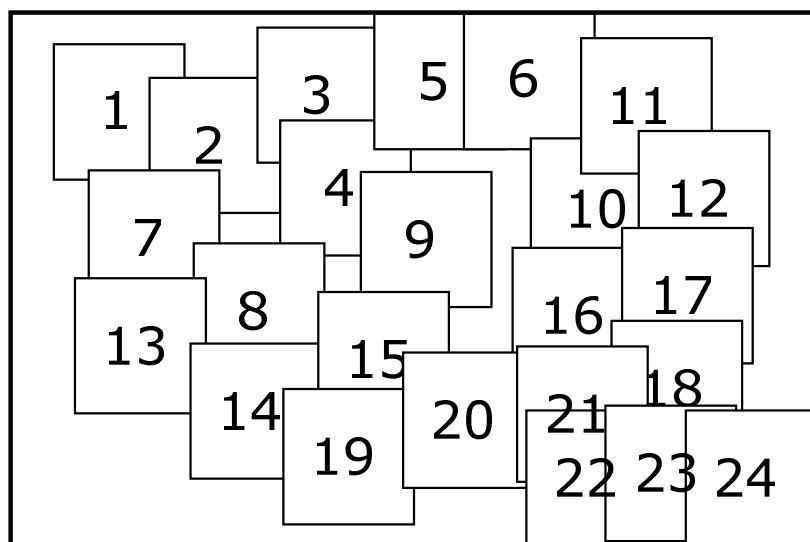
8.3.2. Predykcja z kompensacją ruchu



Predykcja z kompensacją ruchu



obraz odniesienia



aktualnie analizowany obraz

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24

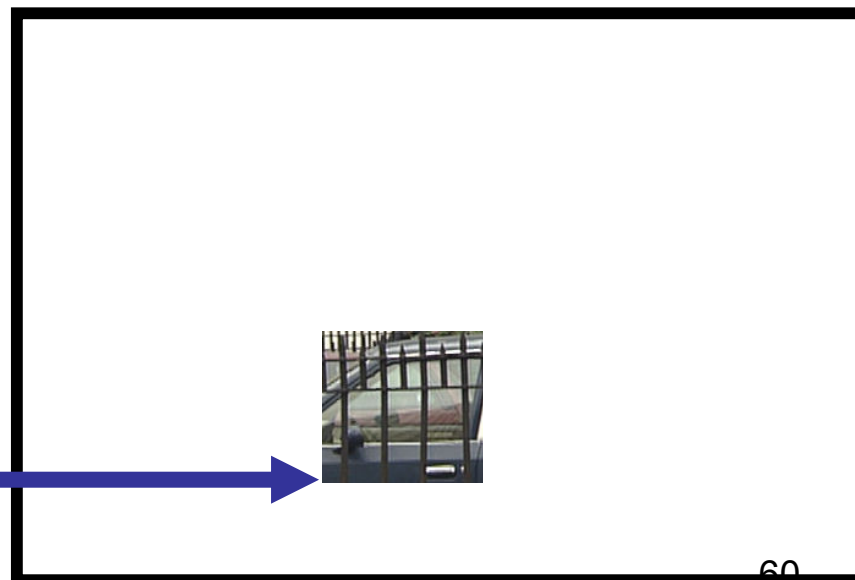
Obraz odniesienia (zdekodowany)



Obraz aktualnie kodowany



**Predykcja obrazu
kodowanego**



Ruch kamery i ruch obiektu !!

Predykcja z kompensacją ruchu

Kodowanie z wykorzystaniem predykcji
z kompensacją ruchu, ale bez przesyłania błędu predykcji



Predykcja z kompensacją ruchu

Kodowanie z wykorzystaniem predykcji
z kompensacją ruchu, ale bez przesyłania błędu predykcji

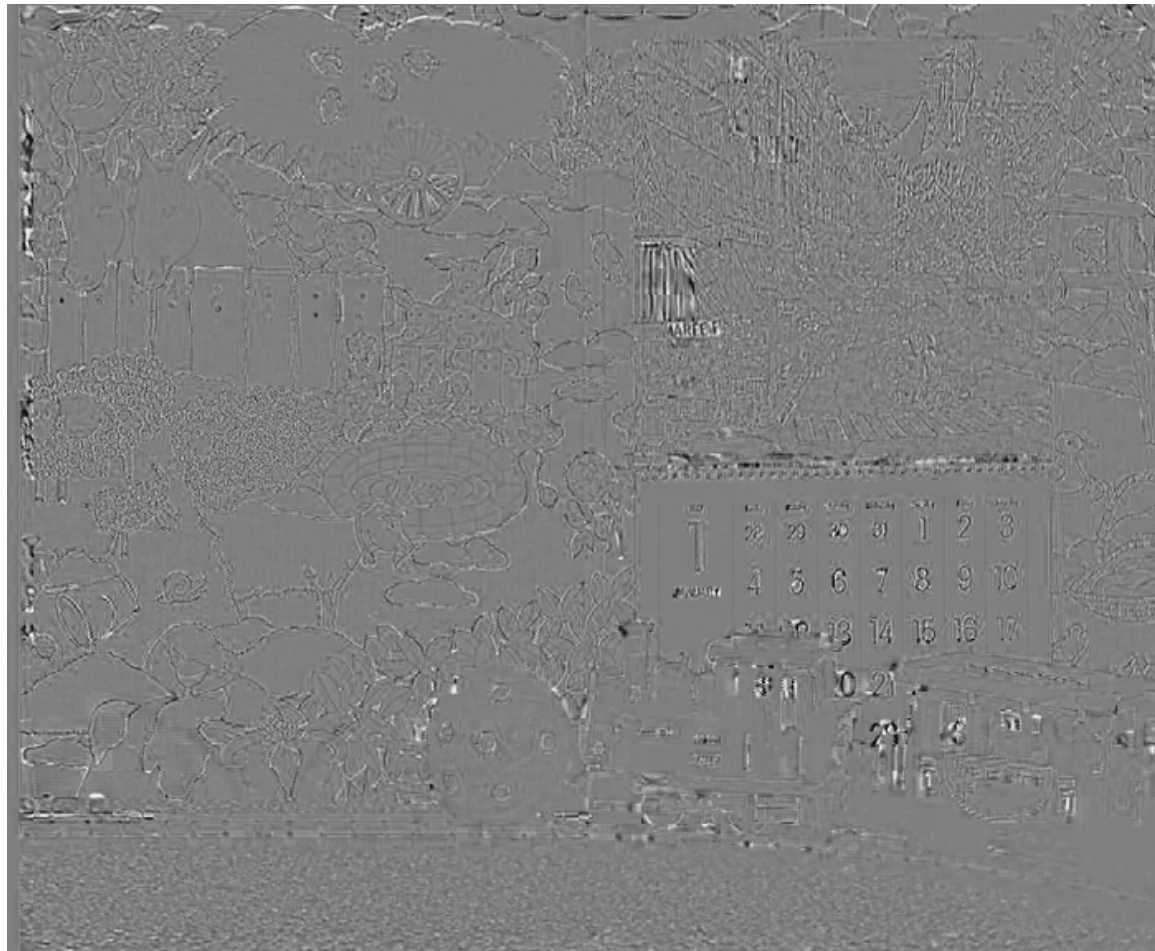


Predykcja z kompensacją ruchu

Kodowanie z wykorzystaniem predykcji
z kompensacją ruchu, ale bez przesyłania błędu predykcji



Błąd predykcji

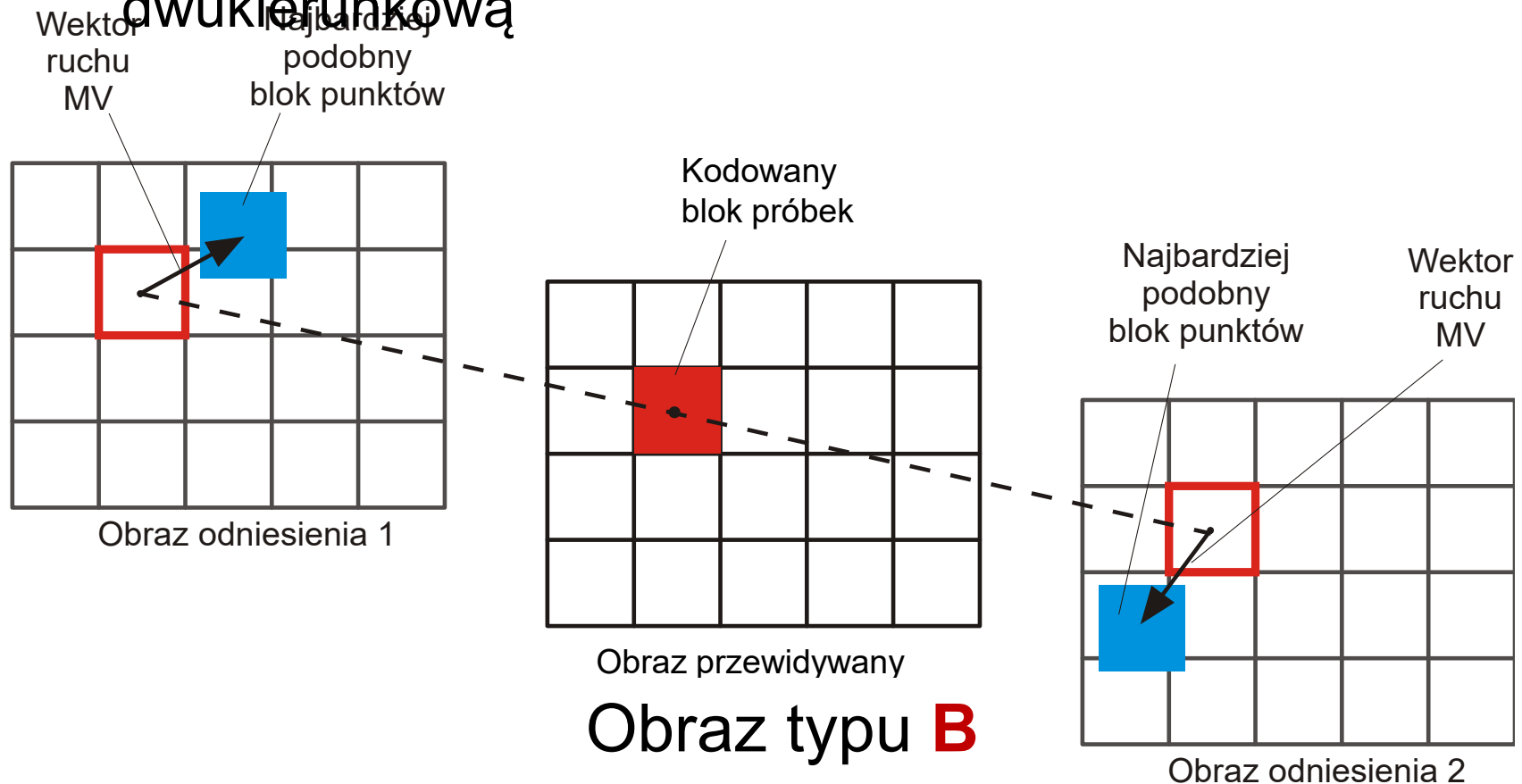


8.3.3. Obrazy typu I, P oraz B

- **Obraz typu I (INTRA):**
 - Kodowany **w całości** w trybie **wewnątrzbrazowym**.
 - Można go zdekodować niezależnie od pozostałych obrazów.
- **Obraz typu P (INTER):**
 - Kodowany **przynajmniej w części** w trybie **międzybrazowym**.
 - Niektóre części obrazu mogą być zakodowane w trybie wewnątrzbrazowym.
 - Zdekodowanie możliwe tylko po zdekodowaniu obrazu odniesienia.

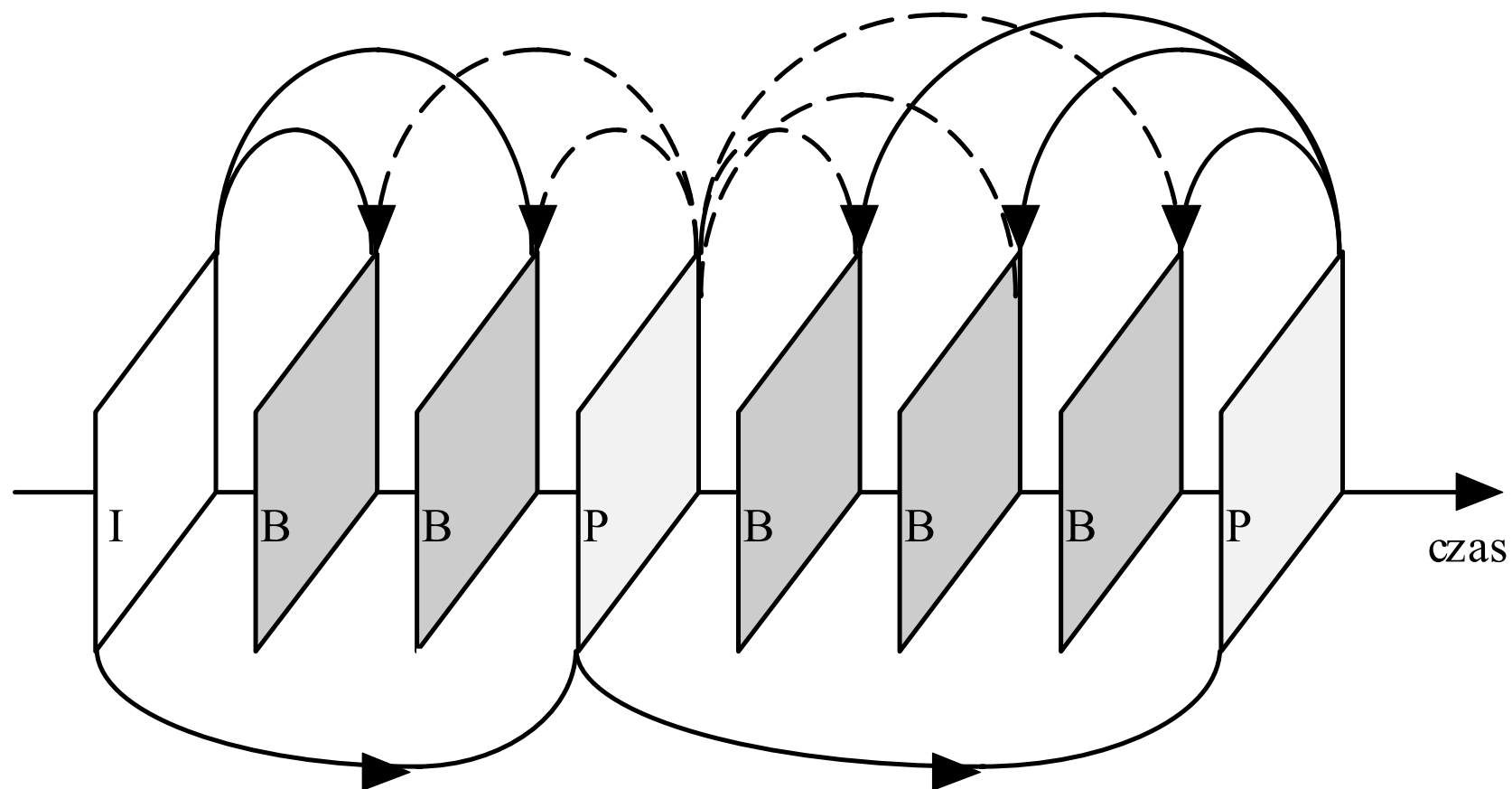
Obraz typu B

Kodowanie międzyobrazowe z predykcją dwukierunkową



- Podwójna estymacja ruchu !!
- Wektor ruchu wskazuje najbardziej podobny blok, niekoniecznie odpowiada ruchowi fizycznemu.

Obrazy typu B (c.d.)



Błąd predykcji

Predykcja jednokierunkowa

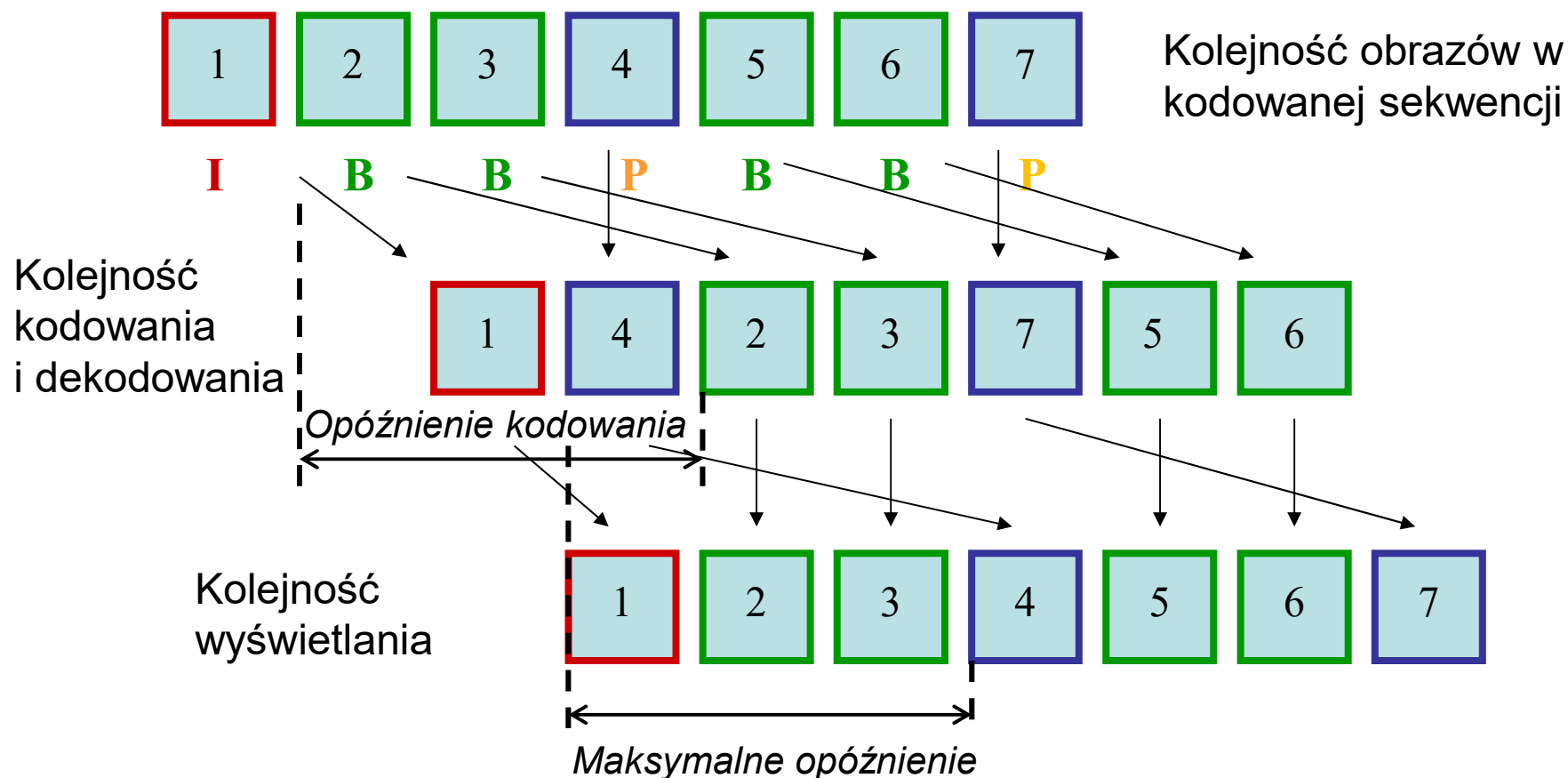


Błąd predykcji

Predykcja dwukierunkowa - obrazy typu B

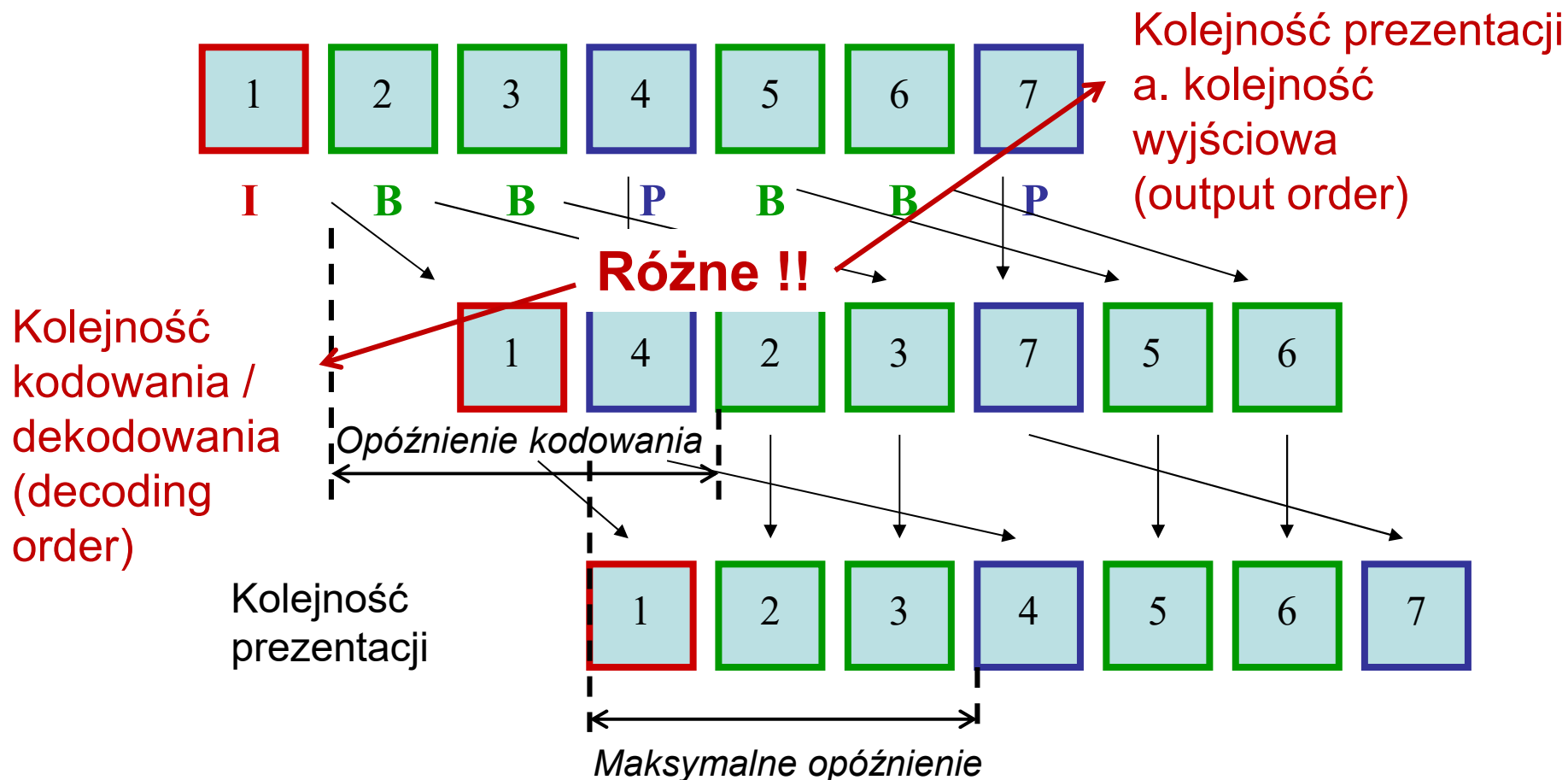


Opóźnienia związane z obrazami typu B



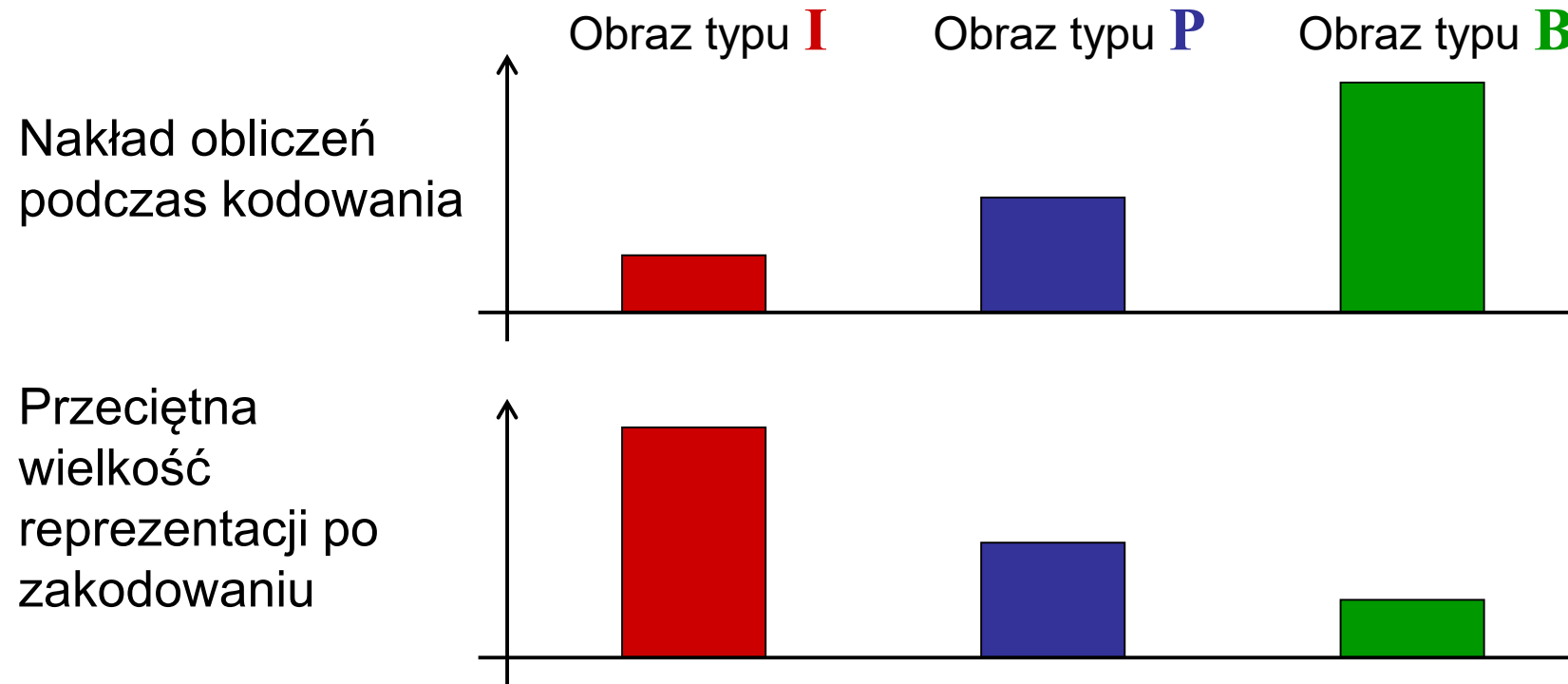
Przykład: Częstotliwość obrazów wynosi 25 Hz (SDTV w Europie)
 - samo zastosowanie 2 obrazów typu B między kolejnymi obrazami typu I/P wprowadza dodatkowe opóźnienie **120 ms**.

Opóźnienia związane z obrazami typu B



Przykład: Częstotliwość obrazów wynosi 25 Hz (SDTV w Europie)
 - samo zastosowanie 2 obrazów typu B między kolejnymi obrazami typu I/P wprowadza dodatkowe opóźnienie **120 ms**.

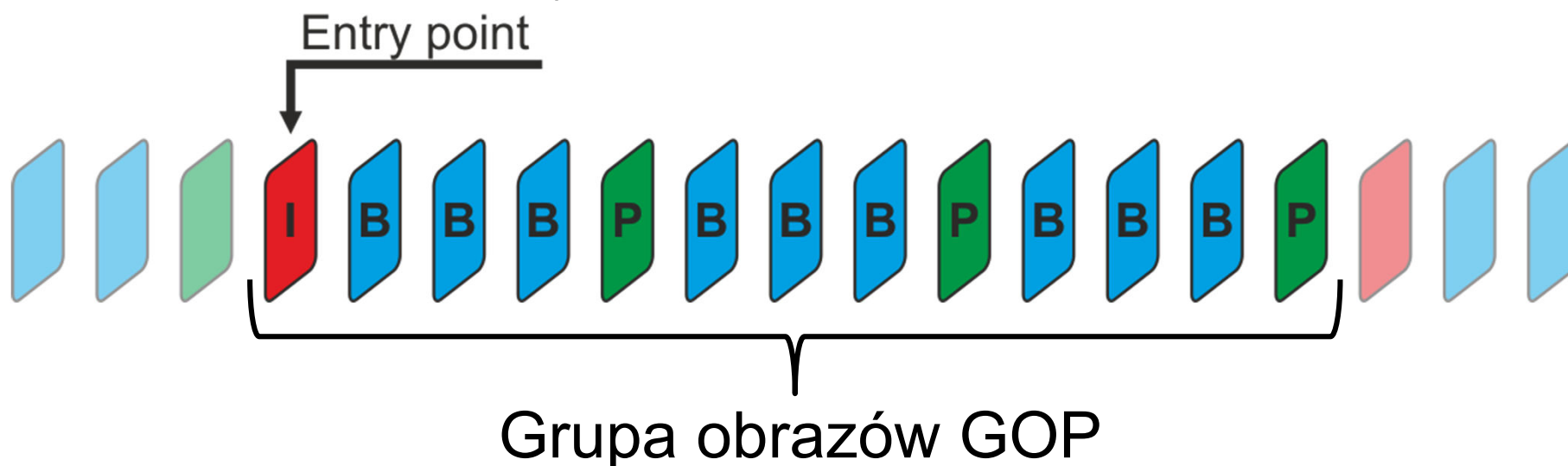
Porównanie różnych typów obrazów



Wysokości słupków są poglądowe
i nie odzwierciedlają wartości liczbowych

8.4. Grupa obrazów *Group of pictures (GOP)*

- GOP – grupa kolejnych obrazów w zakodowanym strumieniu
- Zawiera punkt dostępu (obraz typu I)
- Typy obrazów:
 - Typ I
 - Typ P
 - Typ B



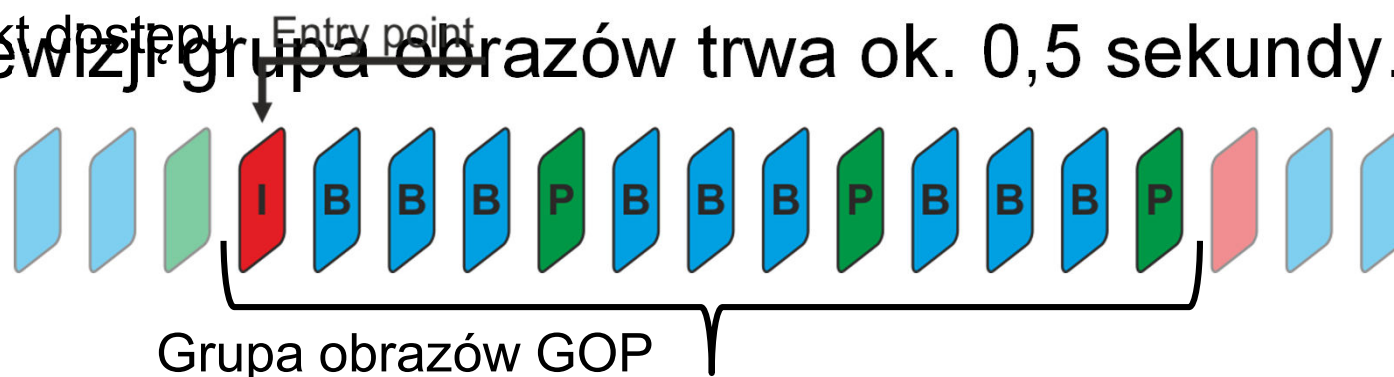
Grupa obrazów *Group of pictures (GOP)*

Wizja jest dekodowalna grupami obrazów.

Dlaczego wydziela się grupy obrazów ?

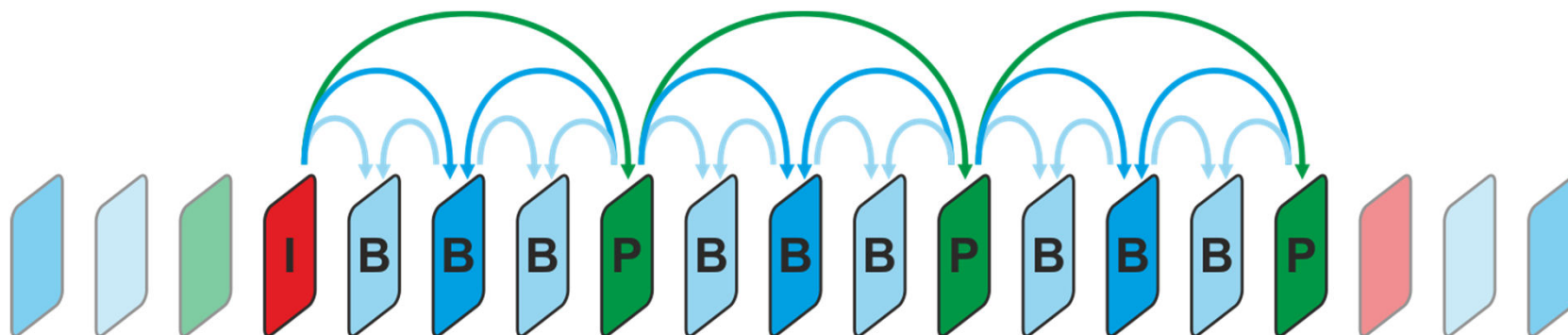
- Możliwość dekodowania wizji od pewnego punktu (późniejsze włączenie odbiornika, zmiana kanałów telewizyjnych/internetowych).
- Szybkie przewijanie wizji z nośnika cyfrowego.
- Ograniczenie propagacji błędów.

W telewizji grupa obrazów trwa ok. 0,5 sekundy.



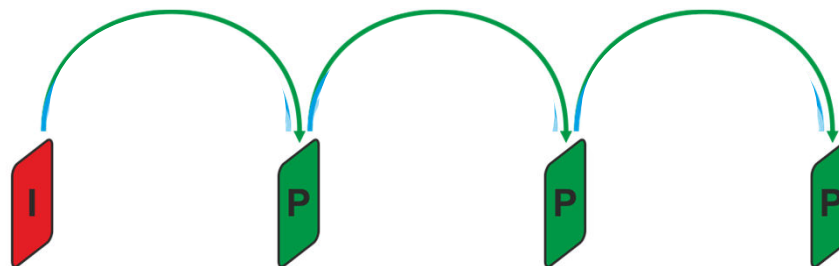
Hierarchiczna grupa obrazów

Lepsza efektywność kompresji + skalowalność czasowa

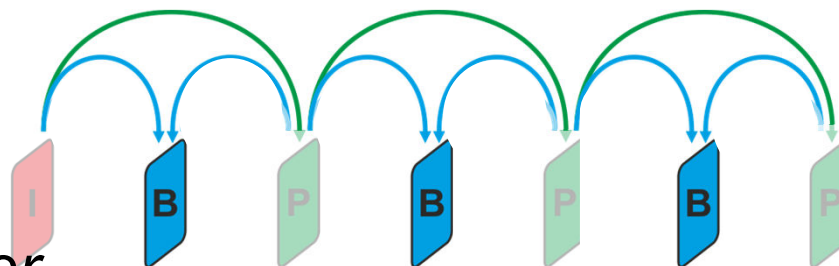


Skalowalność czasowa z wykorzystaniem hierarchicznych obrazów B

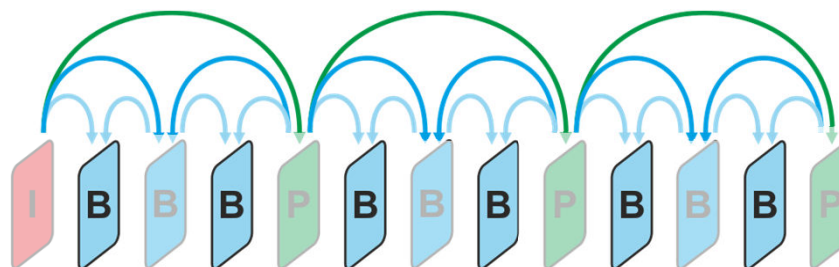
Warstwa podstawowa
Base layer



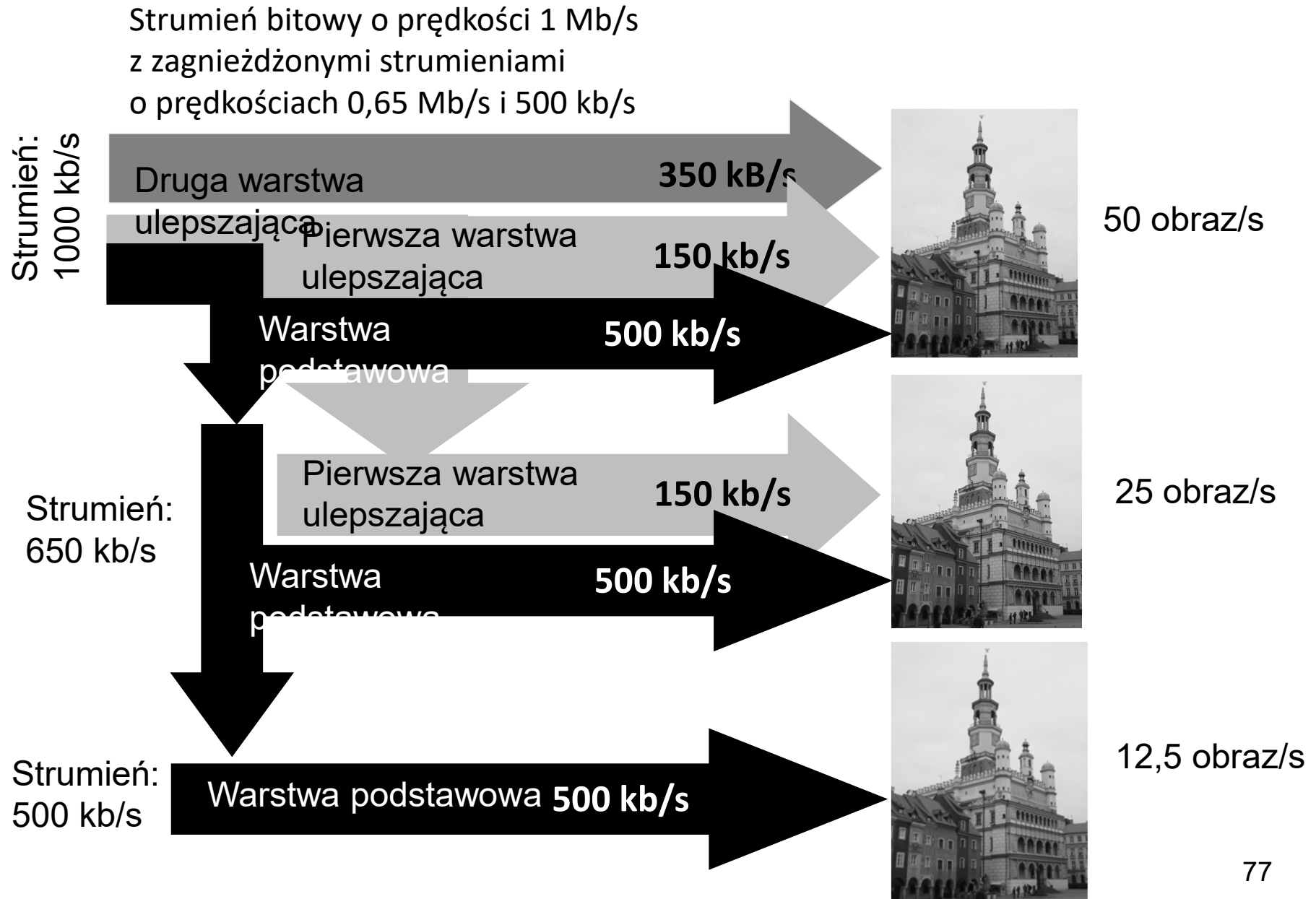
Pierwsza warstwa
ulepszająca
First enhancement layer



Druga warstwa
ulepszająca
Second enhancement layer



Skalowalność czasowa



Rodzaje obrazów I

w kodeku HEVC

- IDR (Instantaneous Decoding Refresh)

Natychmiastowe Odświeżenie Dekodera

Zakodowanie / zdekodowanie obrazu IDR powoduje wykasowanie pamięci obrazów odniesienia.

- CRA (Clean Random Access)
Czysty Dostęp Swobodny

Grupa obrazów rozpoczynająca się obrazem CRA

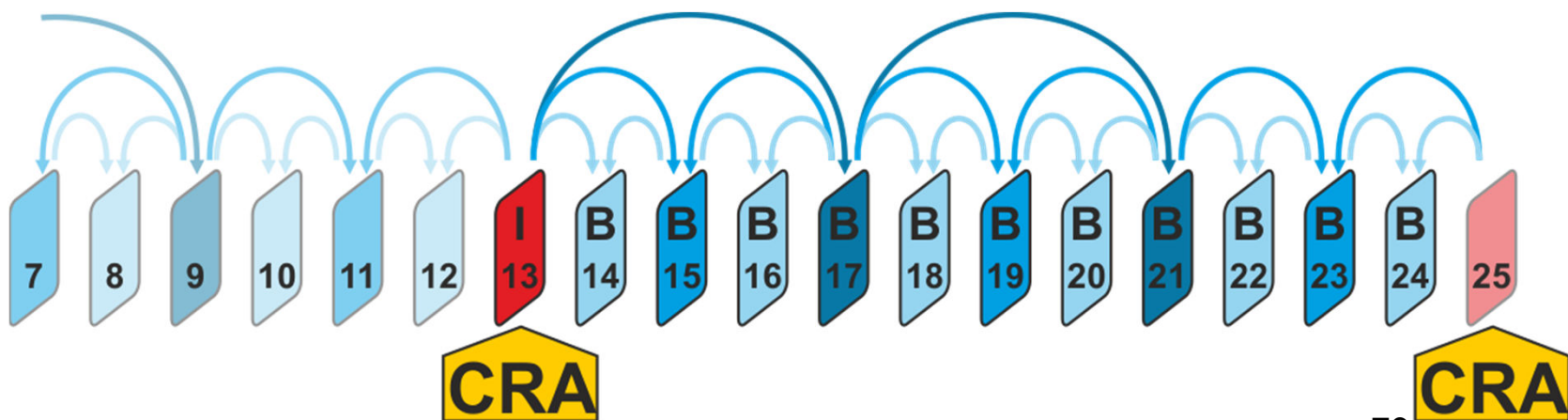
Porządek wyjściowy:



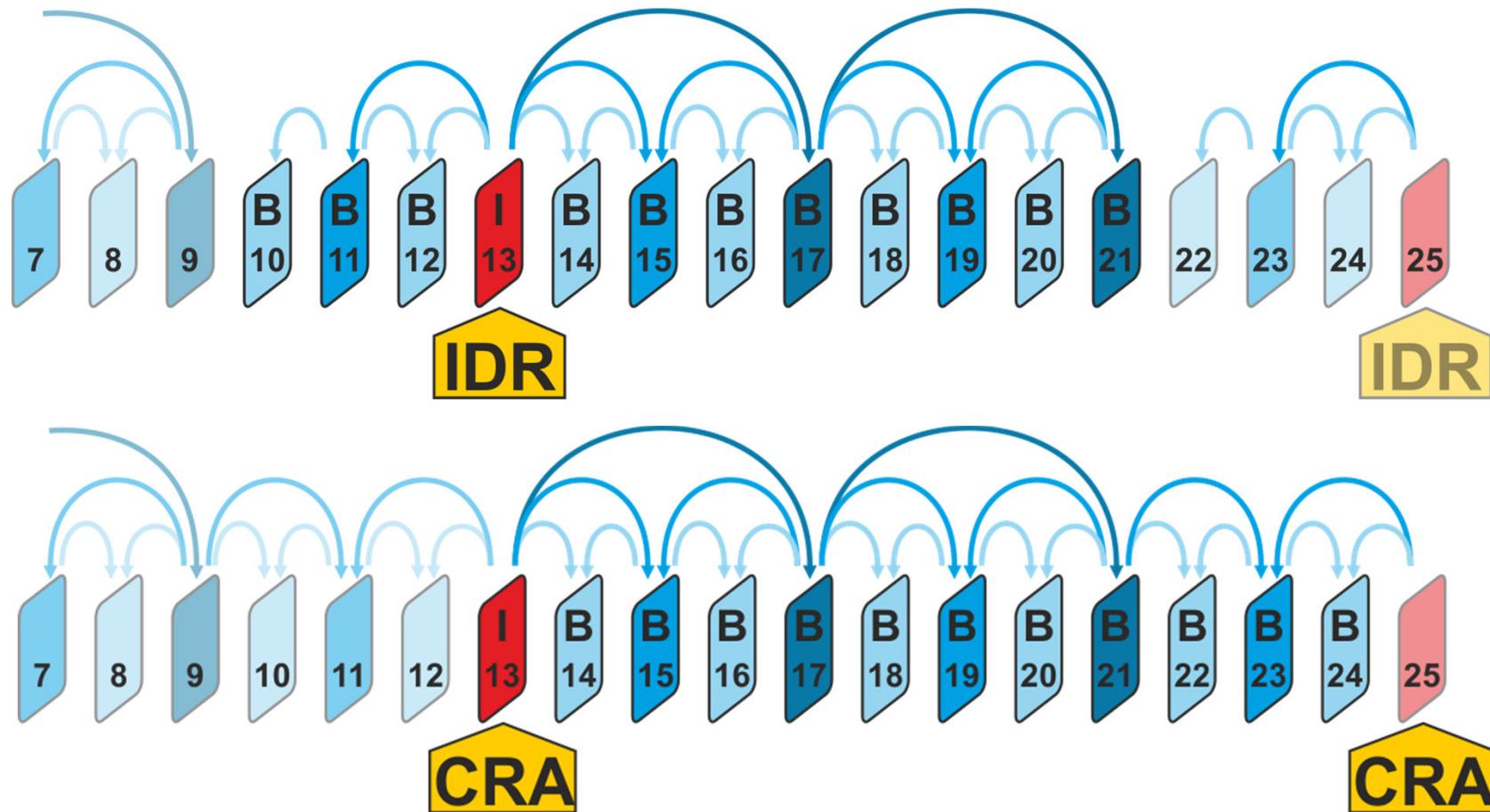
Porządek kodowania:



GOP :



Porównanie IDR oraz CRA



- Użycie CRA poprawia efektywność kompresji (~6%).
- Typowe rozwiązanie (telewizja) :
IDR co 2 sekundy, CRA co 0,5 sekundy.

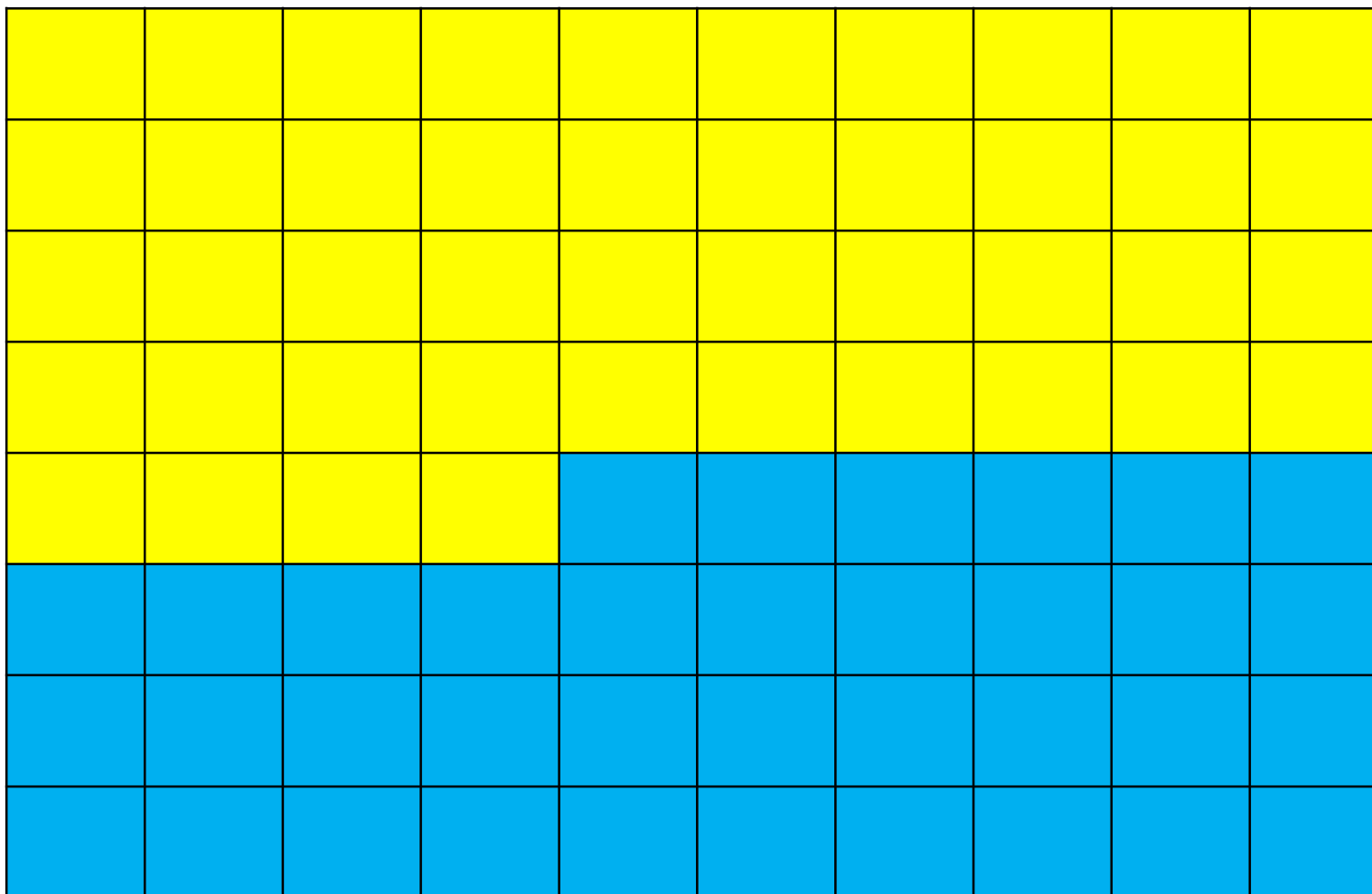
8. 5. Tworzenie i składnia strumienia binarnego

- Plaster (*slice*) jest całym obrazem lub jego częścią.
- Plaster jest najmniejszą strukturą rozpoczynającą się kodem startowym.
- Plaster jest w całości kodowany entropijnie (arytmetycznie) – ciąg bitów dekodowalny tylko w całości.

W JPEG: segment kodowany entropijnie.

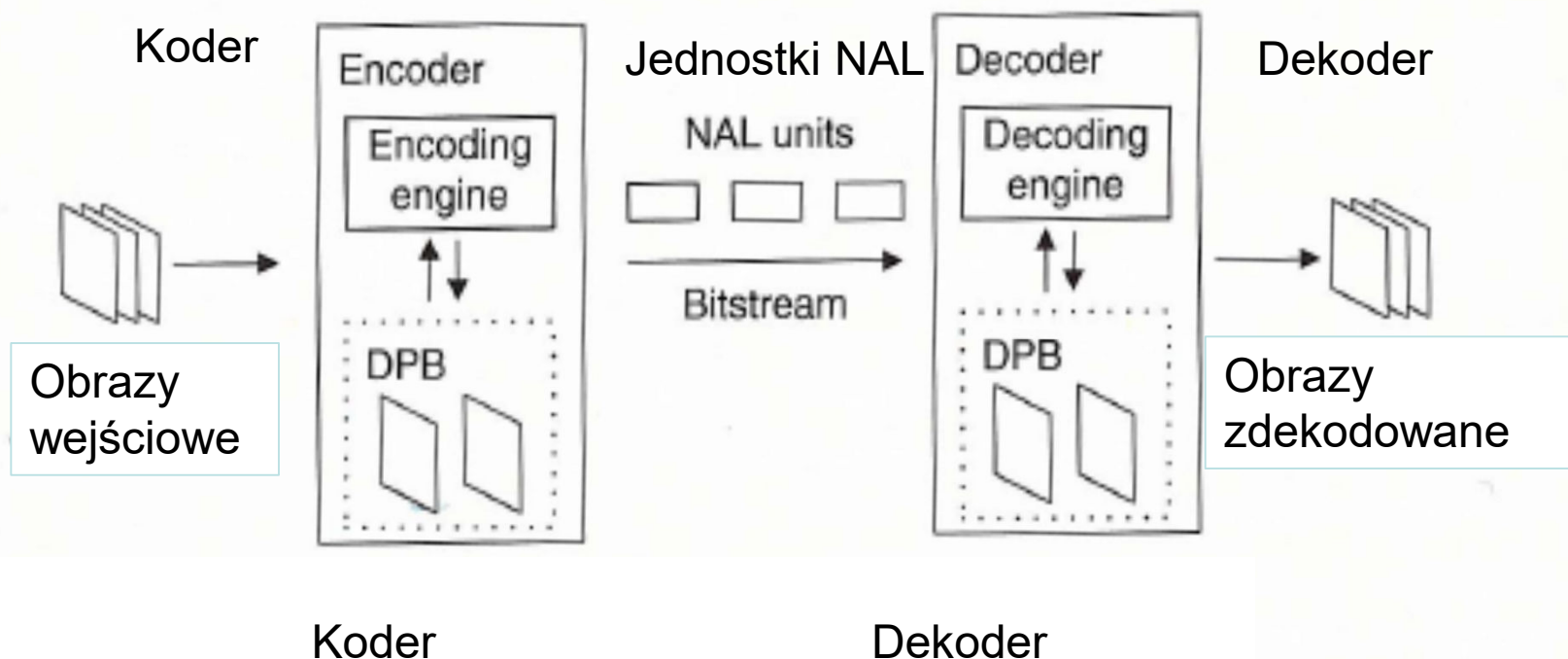
Podział obrazu na plastry

Przykład



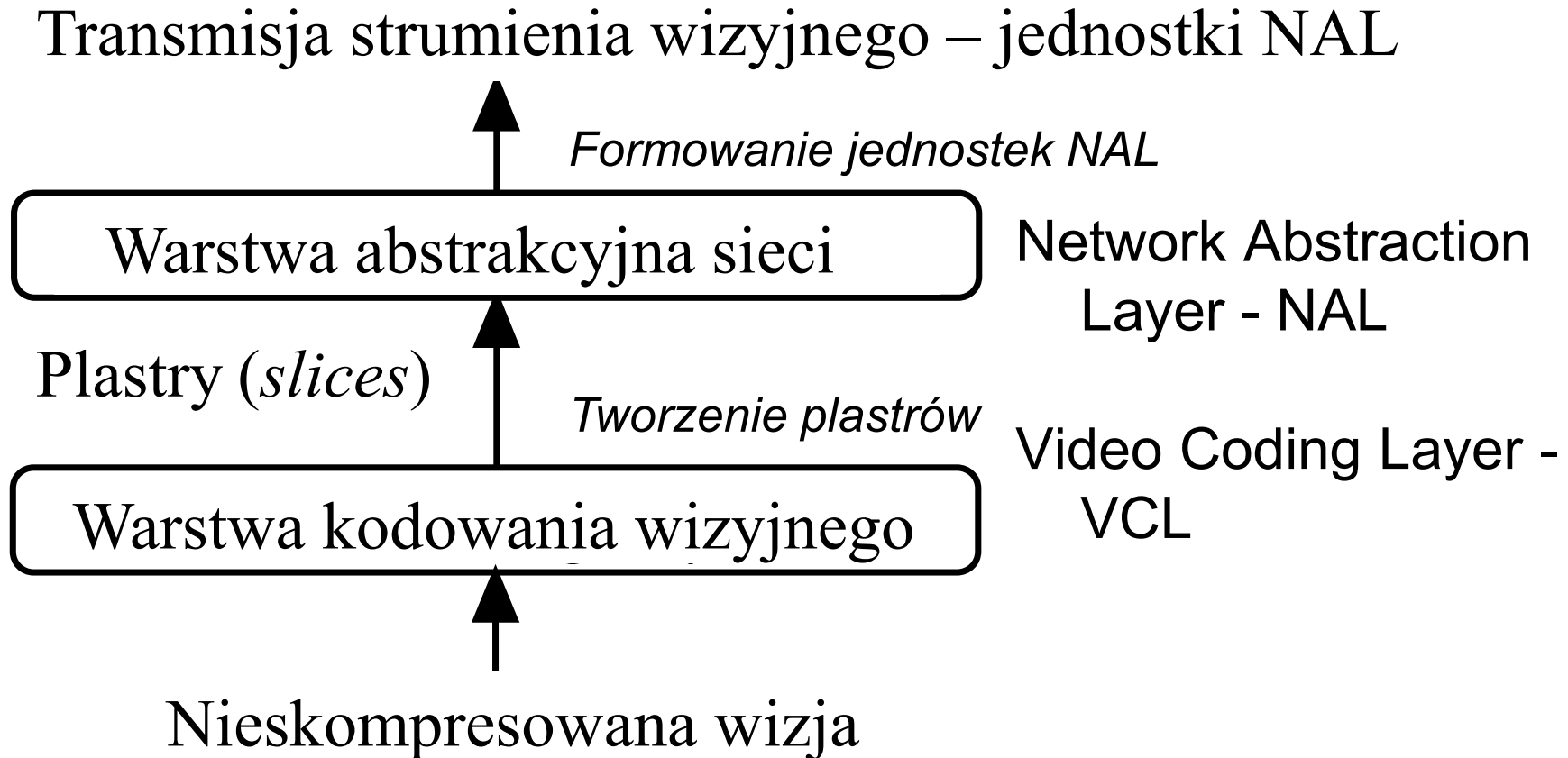
Składnia strumienia – wysokiego poziomu

- Jednostki Warstwy Abstrakcyjnej Sieci – NAL (Network Abstraction Layer) units



DPB – Decoded Picture Buffer – bufor obrazów zdekodowanych

Struktura kodera



Plaster (*slice*) – jednostka danych zakodowanych entropijnie
– zazwyczaj cały obraz, ewentualnie jego część

Jednostki NAL

Jednostki NAL w formacie strumienia bajtowego:

Np. w H.320 lub MPEG-2 TS (H.222)

Przedrostek kodu startowego = start code prefix (3 bajty) : 00 00
01

Jednostki NAL dla systemów przesyłania pakietów

Bez przedrostków

VCL = Video Coding Layer

VCL oraz Non-VCL NAL units

Jednostki Non-VCL NAL zawierają **zbiory parametrów**

Typ jednostki NAL podany jest w nagłówku jednostki NAL (w AVC
: nagłówek to 1 bajt, w HEVC: nagłówek to 2 bajty)

Przedrostek kodu startowego

- 3 bajty: szesnastkowo 00 00 01.
- Unikalny ciąg bitów.
- Służy wyłącznie do zaznaczania miejsca w strumieniu odpowiadającego restartowaniu kodeka entropijnego.
- Po nim następuje kod startowy wskazujący początek pewnej struktury danych.

Warstwa abstrakcyjna sieci *Network Abstraction Layer (NAL)*

NAL units – całkowita liczba bajtów
nagłówek – 1 bajt – określa typ danych
zawartych w NAL

dodatkowo wstawia się: *emulation prevention bytes*

Koder:

001 – emulowany przedrostek kodu startowego

0031 – emulacja usunięta

003 → 0033

Dekoder:

003 → 00

Strumień binarny

- Strumień jednostek NAL

Warstwa abstrakcyjna sieci

Jednostki NAL non-VCL → Strumień binarny



Jednostki NAL VCL

Wykrywanie i usuwanie emulacji
przedrostków kodów startowych



Plastry (*slices*)

Wizja

Kodowanie wizji

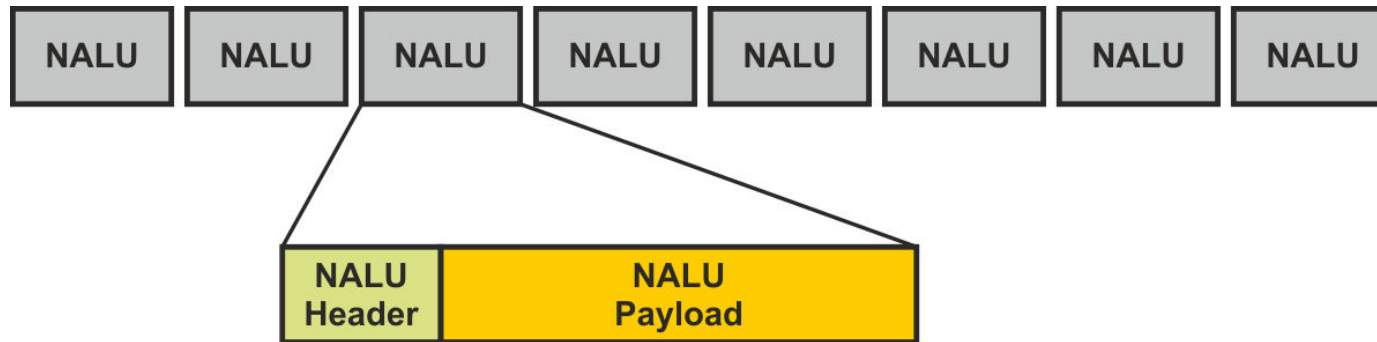


Warstwa kodowania wizyjnego

Składnia wysokiego poziomu

Dwa typy jednostek NAL :

- 1) VCL (Video Coding Layer)
- 2) Non-VCL

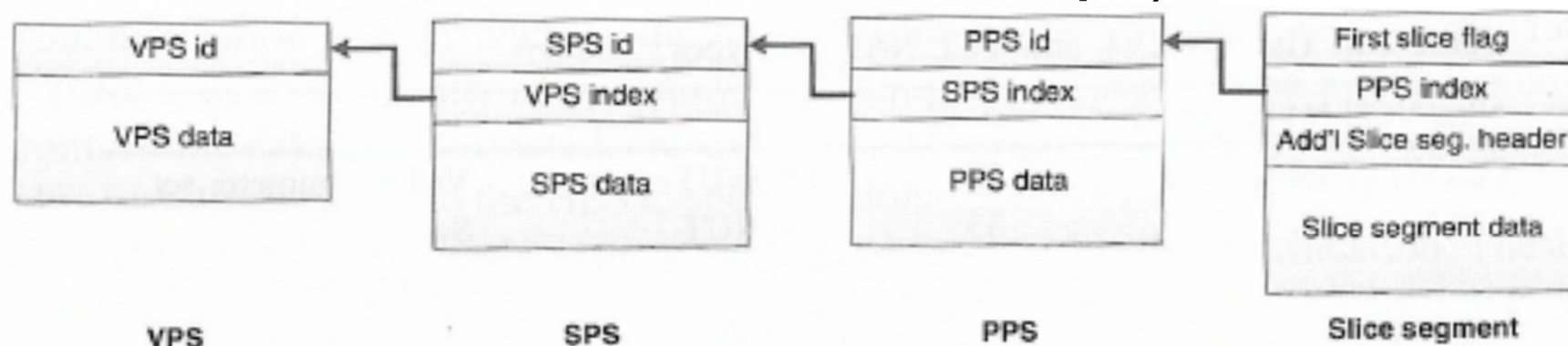


Nagłówek NALU - 2 bajty (HEVC), 1 bajt (AVC):

- 1 bit Wskaźnik – zawsze "0",
- 6 bits Typ jednostki NAL,
- 6 bits LayerId (identyfikator warstwy np. w kodowaniu skalowalnym)
- 3 bits TID (identyfikator „pod-warstwy” w skalowalności czasowej)

Jednostki Non-VCL NAL

- **VPS** video parameter set – zbiór parametrów wizji
- **SPS** sequence parameter set – zbiór parametrów sekwencji



Plaster (Slice): jeden (najczęściej) lub kilka segmentów plastra.

Segment plastra (Slice segment): niezależnie kodowany entropijnie.

W najprostszym przypadku : **Obraz = 1 plaster = 1 segment plastra.**

Typy jednostek NAL



Slice = plaster

- Non-VCL NAL Units – Jednostki NAL typu „Non-VCL” :
 - Video Parameter Set (VPS) – Zbiór parametrów wizji
 - Sequence Parameter Set (SPS) – Zbiór parametrów sekwencji
 - Picture Parameter Set (PPS) – Zbiór parametrów obrazu
 - Supplementary Enhancement Information (SEI)
 - Dodatkowa Informacja Rozszerzająca
- VCL NAL Units – Jednostki NAL typu „VCL” :
 - Slice of encoded picture – Plaster kodowanego obrazu

Zbiór parametrów wizji – Video Parameter Set



- Aktywowany na początku sekwencji (sekwencja zaczyna się obrazem IDR).
- Nie występuje w AVC.
- Zawiera podstawowe informacje o całej sekwencji (łącznie dla wszystkich warstw):
 - Charakterystyka warstw w sekwencji (np. dla kodowania skalowalnego);
 - Informacje synchronizacyjne (np. jednostki czasu);
 - Parametry obrazów o dużym zakresie dynamicznym (HDR) .

Zbiór parametrów sekwencji – Sequence Parameter Set



- Sekwencja zaczyna się obrazem IDR.
- Sekwencja składa się z jednostek dostępu rozpoczynających się obrazami I (np. CRA).
- SPS zawiera podstawowe informacje o całej sekwencji:
 - Rozmiary ramek lumy, typ próbkowania chromy (np. 4:2:0);
 - Liczby bitów próbek lumy i chromy;
 - Profil, warstwa (tier) i poziom kodeka;
 - Parametry narzędzi kompresji (opcjonalnie);
 - Itp.

Zbiór parametrów obrazu – Picture Parameter Set



- Zawiera:
 - Parametry podziału obrazów;
 - Wstępne określenie kroku kwantowania współczynników transformaty;
 - Informacja o ewentualnym rozszerzeniu nagłówka plastra;
 - Parametry filtrów obrazu odniesienia;
 - Opcjonalne dodatkowe parametry dla niektórych profili (np. „screen content – zawartość ekranu” – dla kodowania również grafiki; np. „range extension – rozszerzony zakres – dla kodowania np. próbek o dużych liczbach bitów).
 - Itp.

Dodatkowa informacja rozszerzająca – Supplemental Enhancement Information SEI

- Typ jednostek NAL Non-VCL.
- Informacje niepotrzebne dekodderowi do zdekodowania wartości próbek.
- Informacje pożyteczne lub konieczne dla prawidłowej prezentacji zdekodowanych obrazów.
- Przykłady:
Definicje typu układu RGB;
Parametry transmitancji opto-
elektronicznej;
Informacje o sposobie umieszczenia w

8.6. Warstwa kodowania wizyjnego

Na przykładzie kodeka HEVC

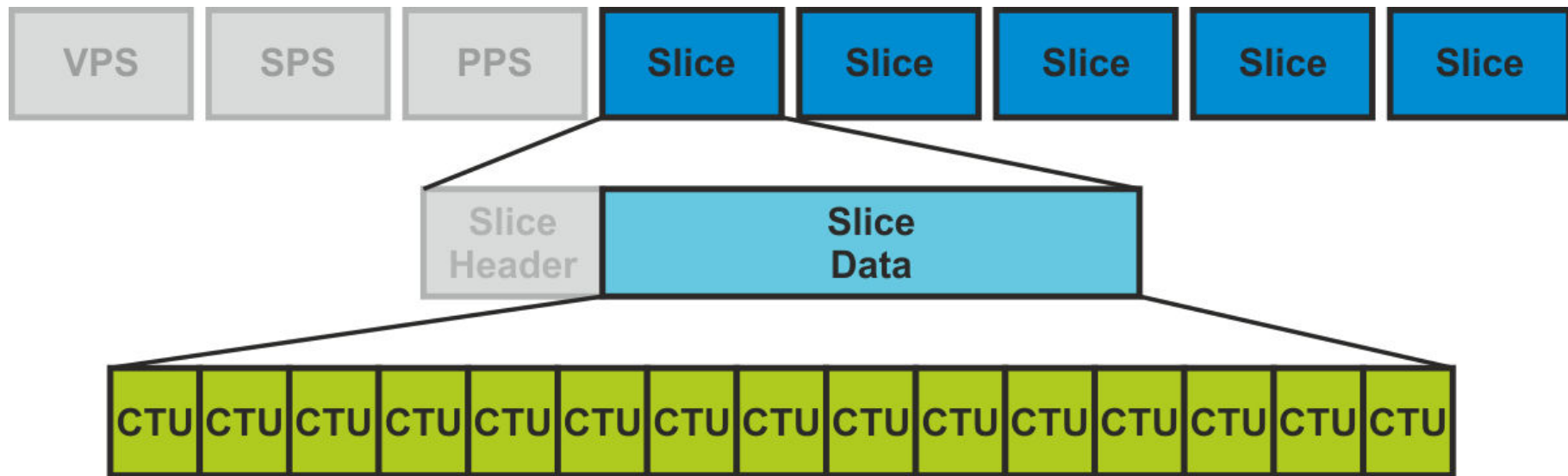
Dane plastrów są kodowane entropijnie
- kodowanie arytmetyczne

Podział obrazów

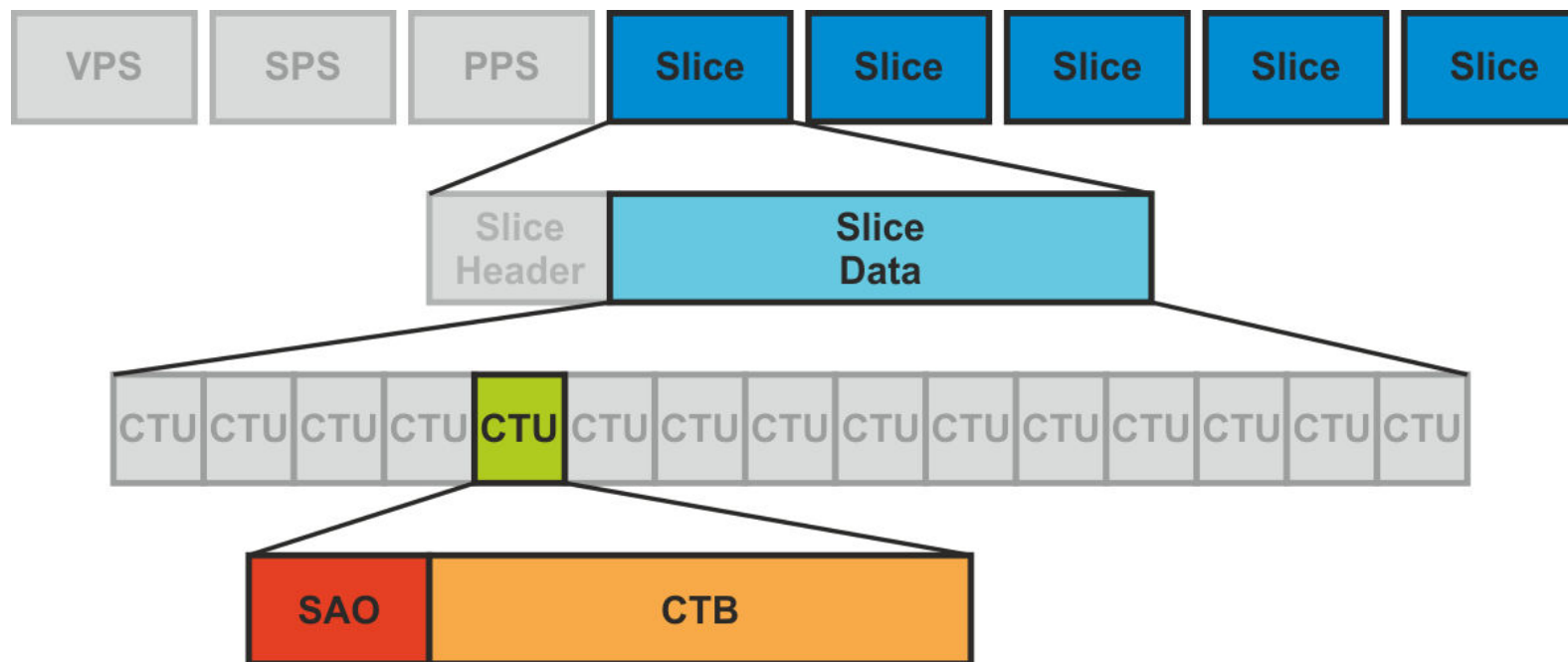
- Obraz składa się z Jednostek Drzewa Kodowania (Coding Tree Units) – HEVC, VVC.
- Rozmiar CTU – **64x64**, 32x32 or 16x16 próbek lumy (HEVC) – rozmiar definiowany jest w SPS.

CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU
CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU
CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU
CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU
CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU
CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU	CTU

Dane plastra

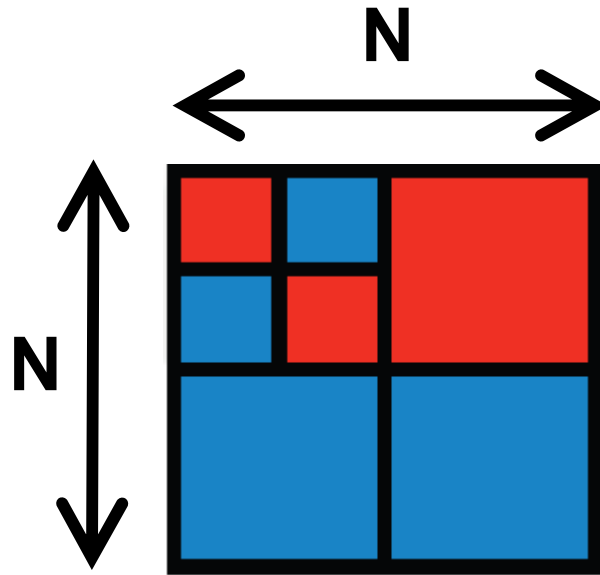


Podział Jednostek Drzewa Kodowania (CTU)



- Sample Adaptive Offset (SAO) – Parametry dla adaptacyjnej zmiany wartości próbek – SAO to jedno z narzędzi usuwania efektów silnej kompresji z obrazu odniesienia.
- CTU składa się z 3 bloków CTB (Coding Tree Block): jednego dla lumy i dwóch dla chrom.

Podział Jednostek Drzewa Kodowania (CTU)

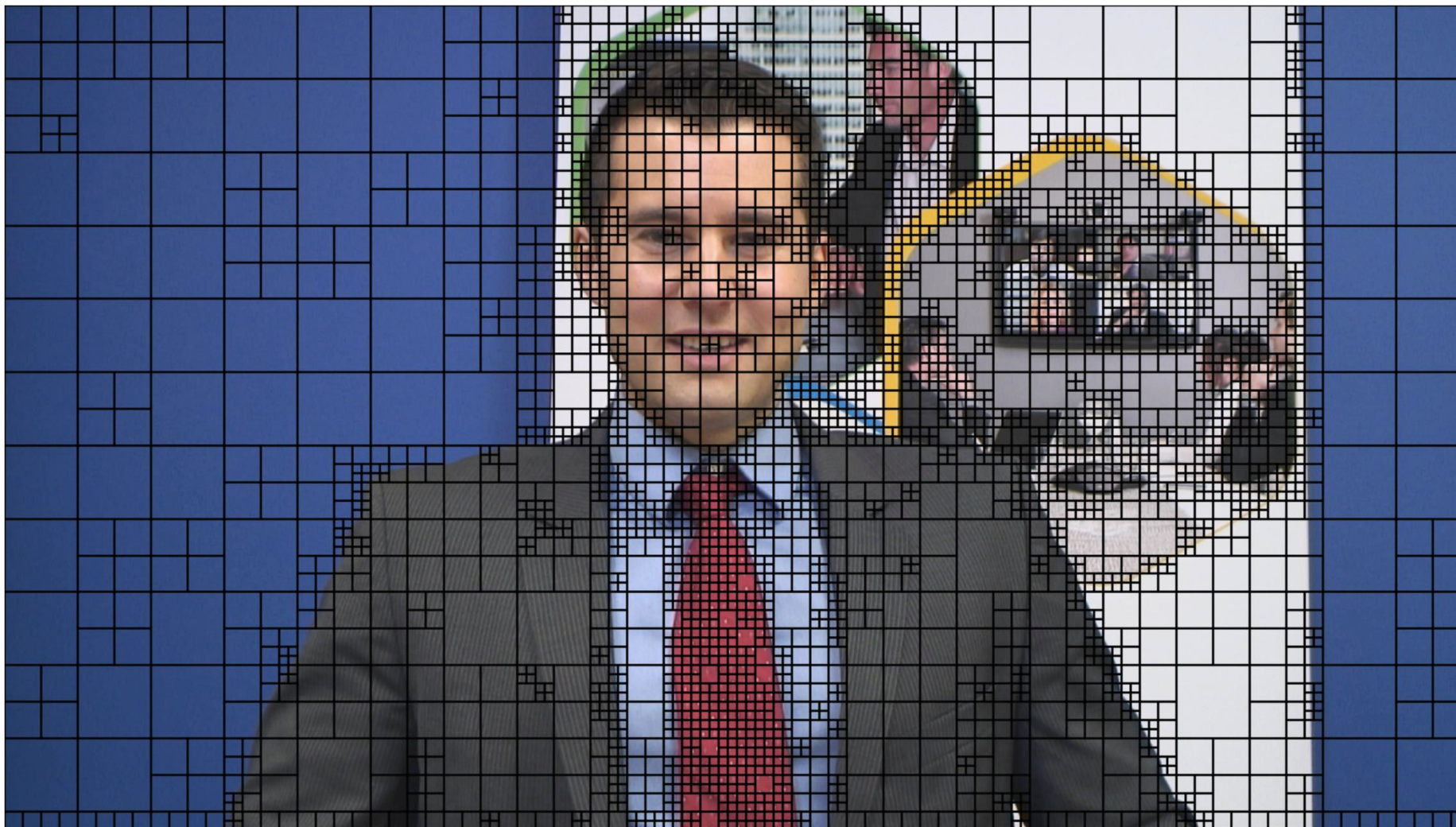


Jednostki Kodowania
Coding Units CUs

$N=16, 32, \text{ or } 64$

Jednostki Kodowania *Coding Units (CUs)* mogą być typu wewnątrzbrazowego (*Intra*) i międzyobrazowego (*Inter*).
Obejmują próbki lummy i chromy
– max. 64×64 do min. 8×8 próbek lummy.

Jednostki kodowania

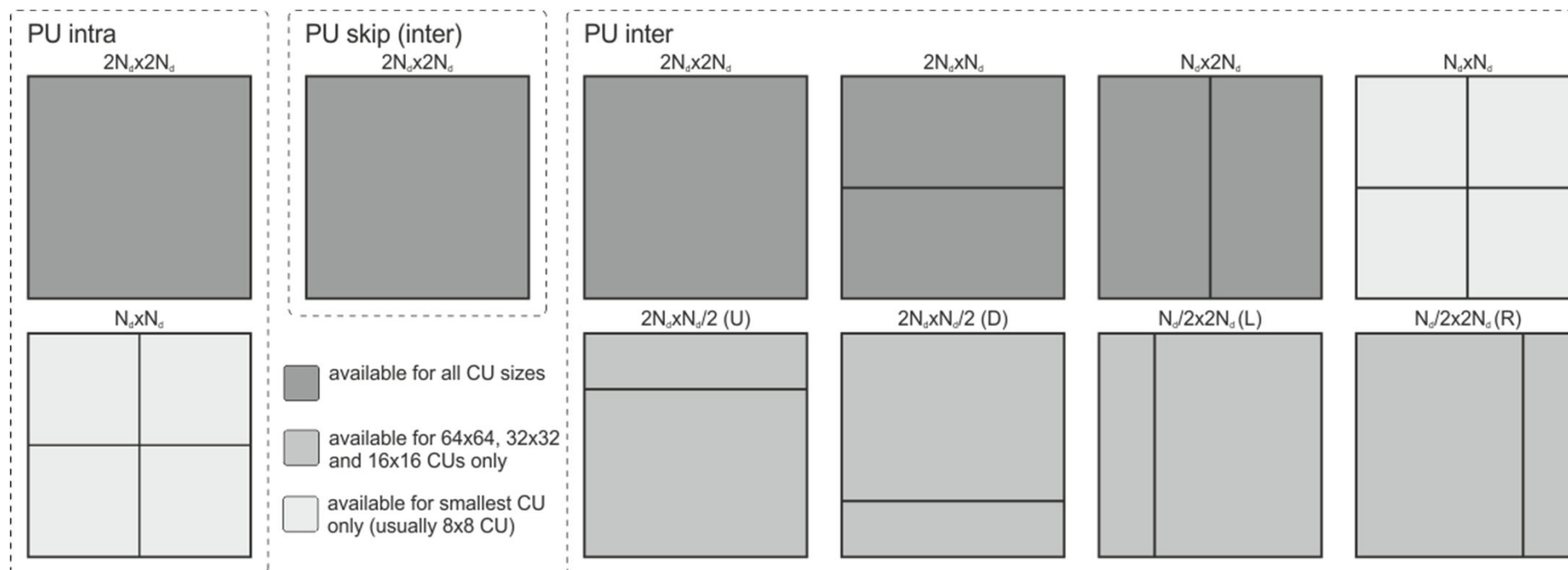


Jednostki kodowania typu SKIP

- **Jednostki kodowane międzyobrazowo.**
- **Nie przesyła się żadnych danych.**
- **Najczęstszy typ jednostek CU – przy przeciętnych prędkościach bitowych nawet 60-90% powierzchni obrazu.**

Podział jednostek kodowania na jednostki predykcji

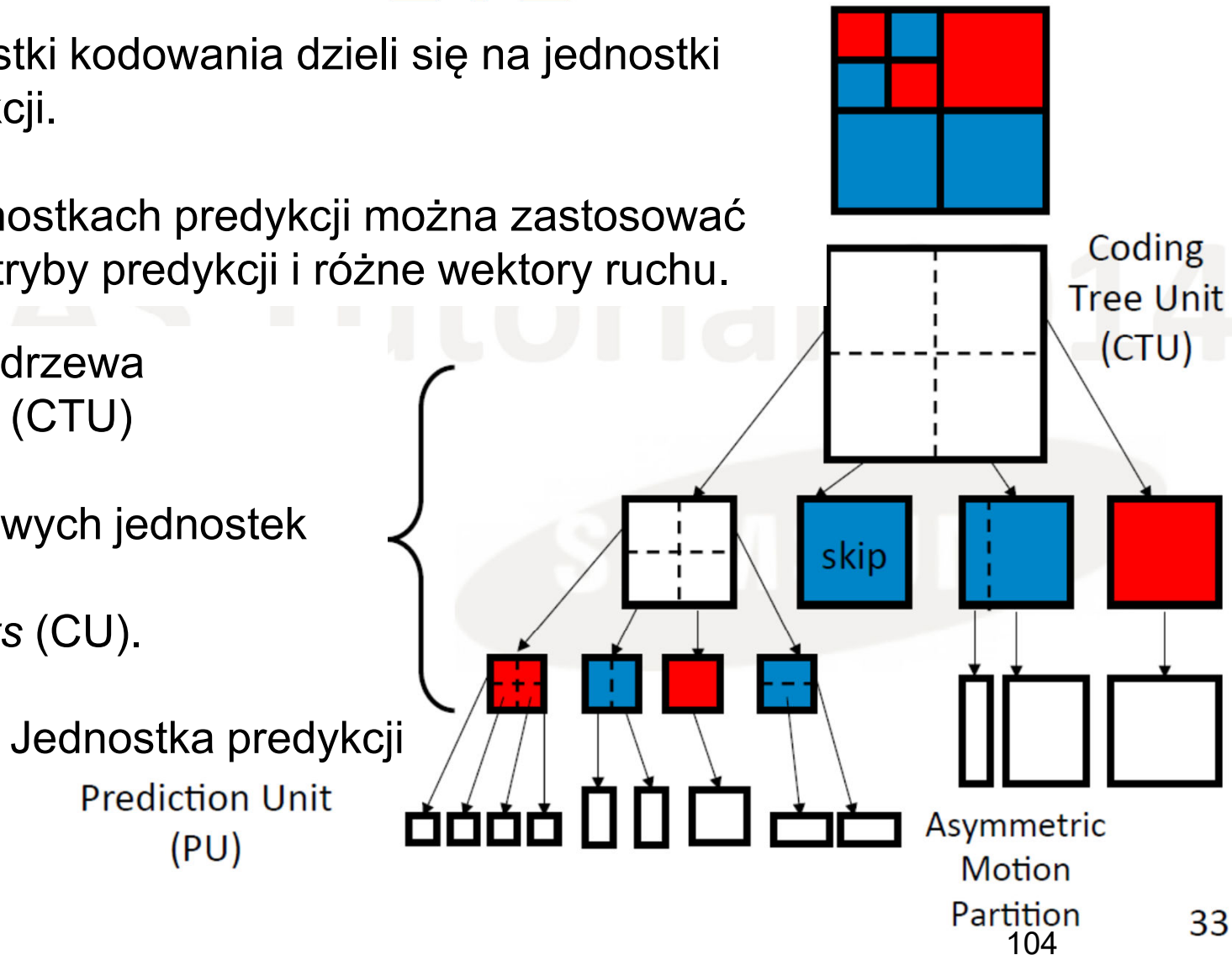
- Jednostki predykcji intra/inter.



Główne jednostki podziału CTU

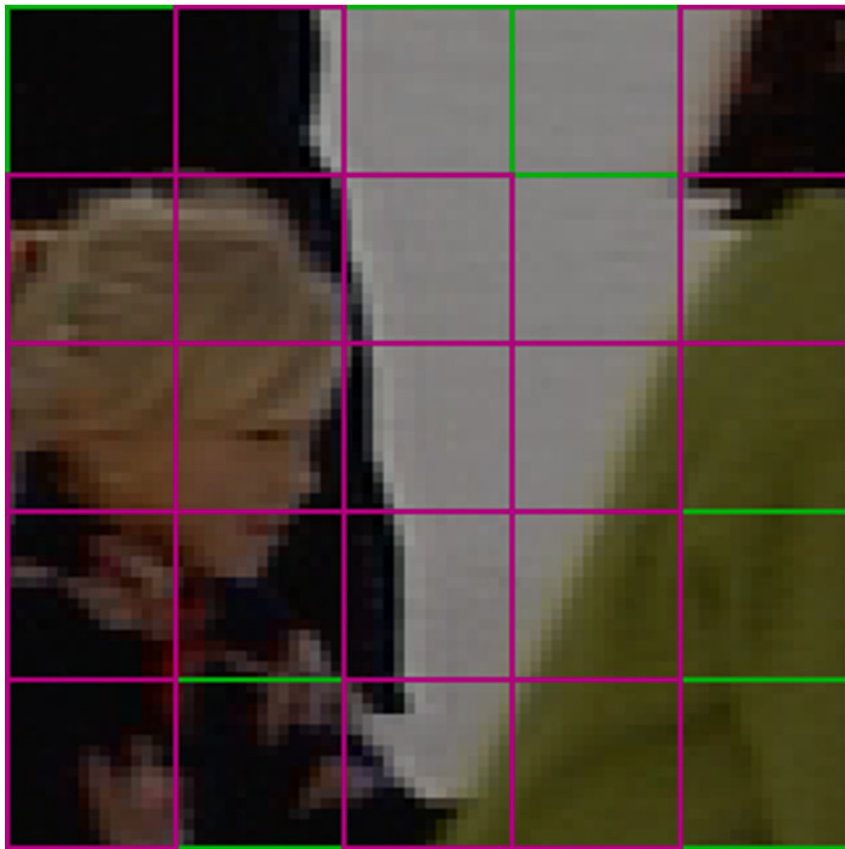
- Jednostki kodowania dzieli się na jednostki predykcji.
- W jednostkach predykcji można zastosować różne tryby predykcji i różne wektory ruchu.

Jednostka drzewa kodowania (CTU) składa się z kwadratowych jednostek kodowania *coding units* (CU).



Predykcja blokami o zmiennych rozmiarach

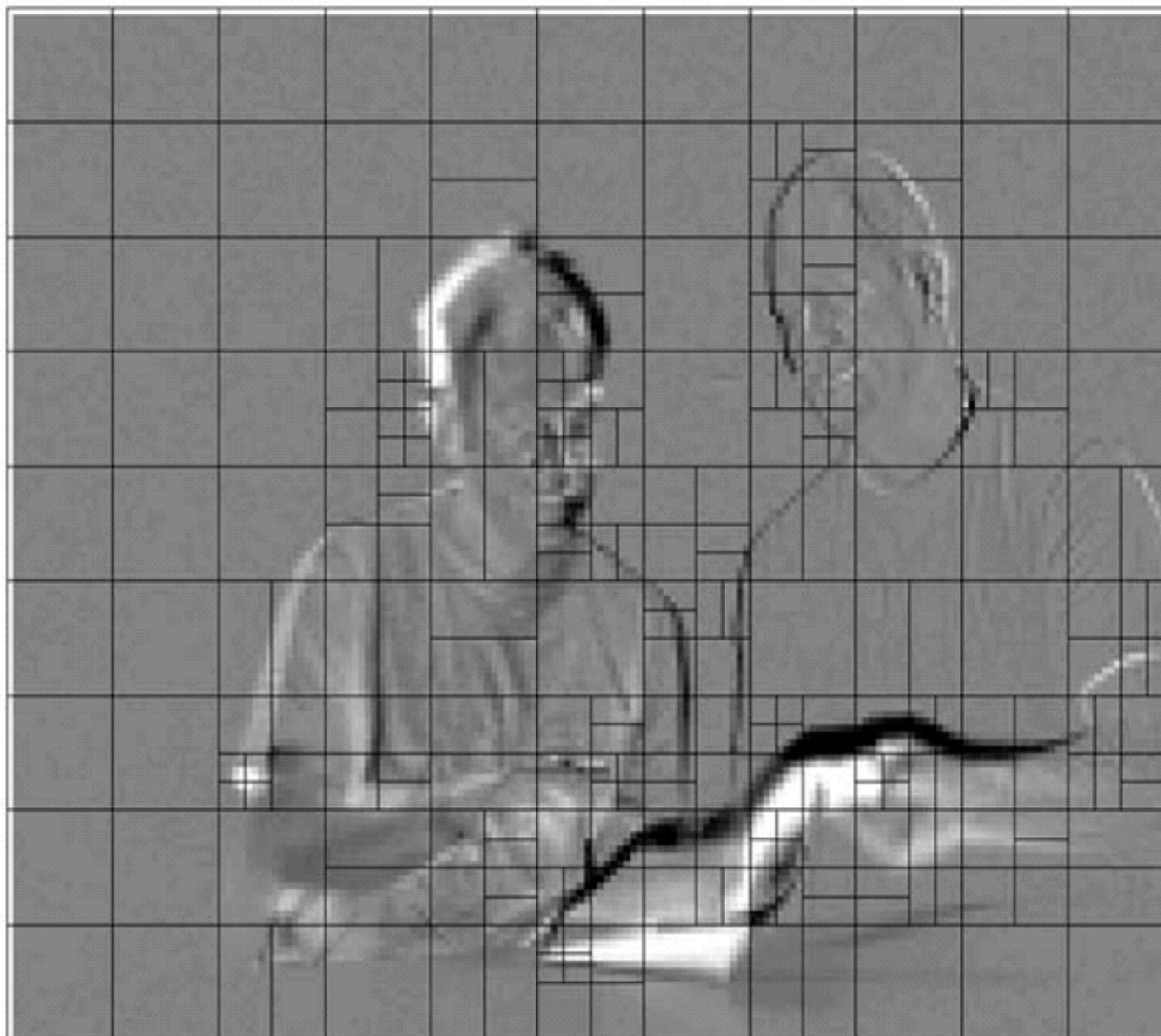
Stały rozmiar bloku



HEVC

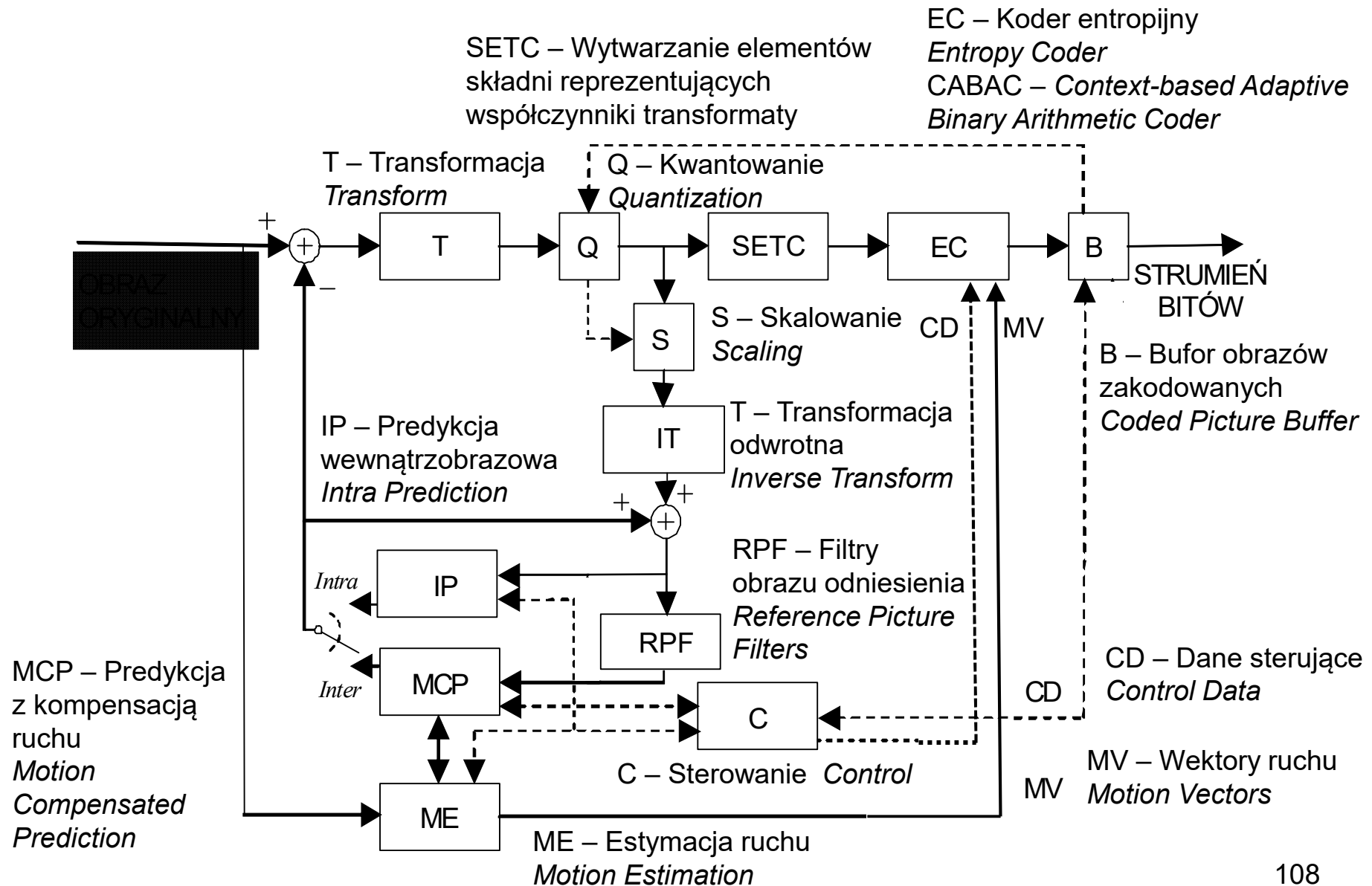


Predykcja w blokach o zmiennych rozmiarach



8.7. Struktura kodeka

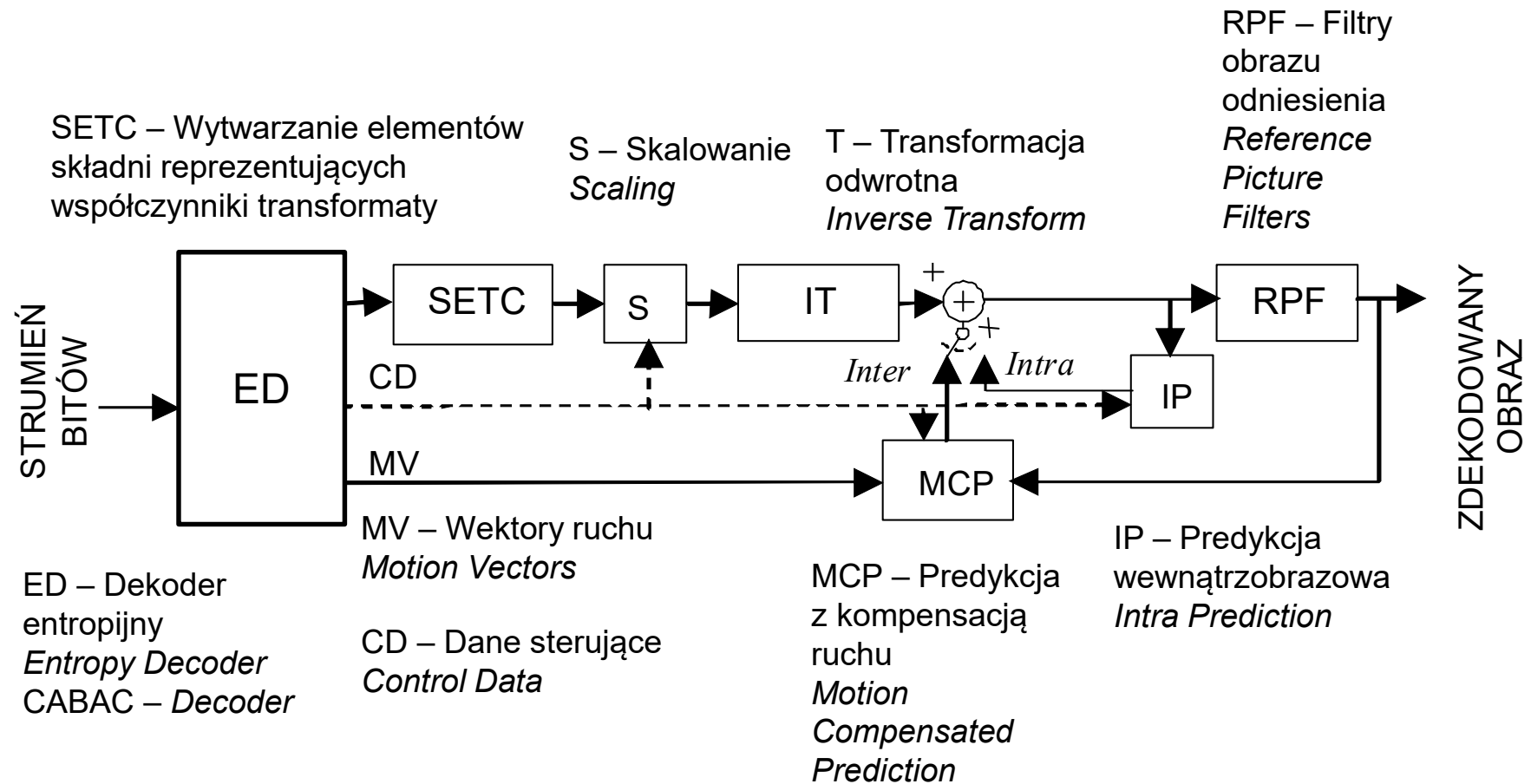
Koder HEVC



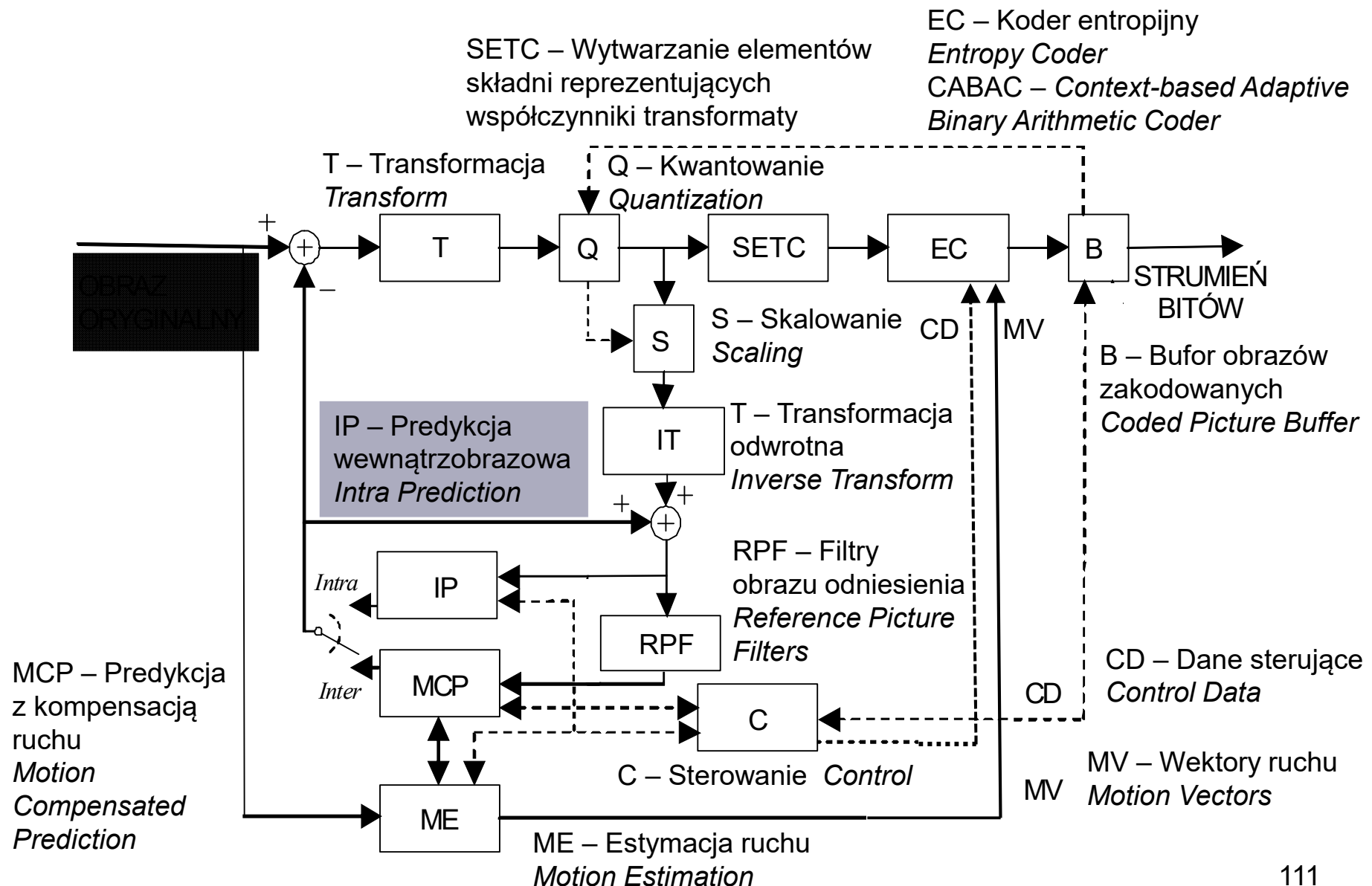
Elementy składowe Strumień bitów



Dekoder HEVC



8.8. Predykcja wewnątrzbrazowa

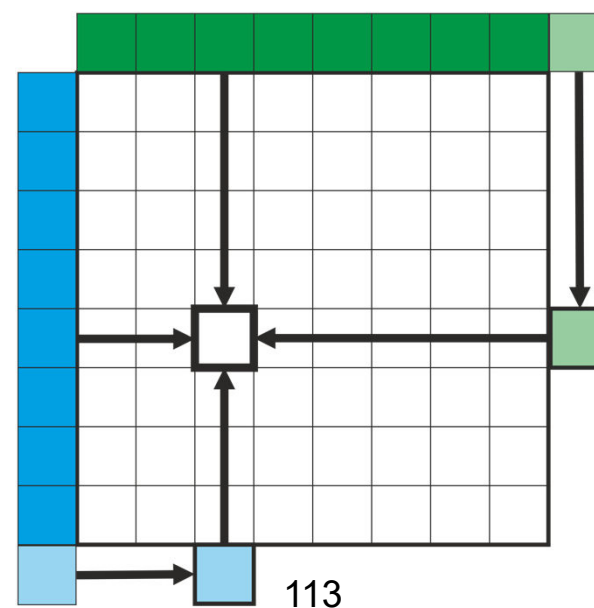
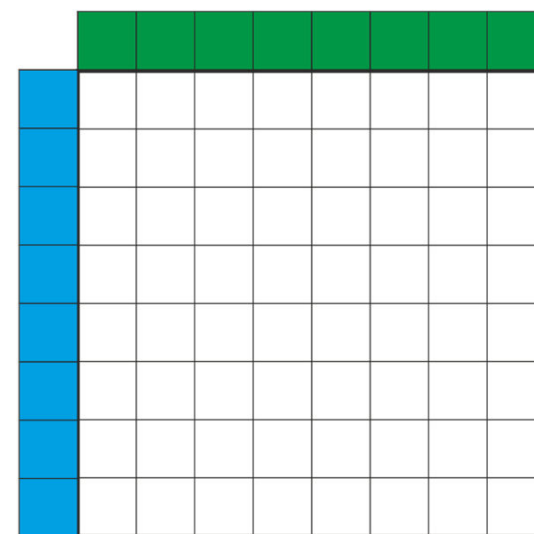
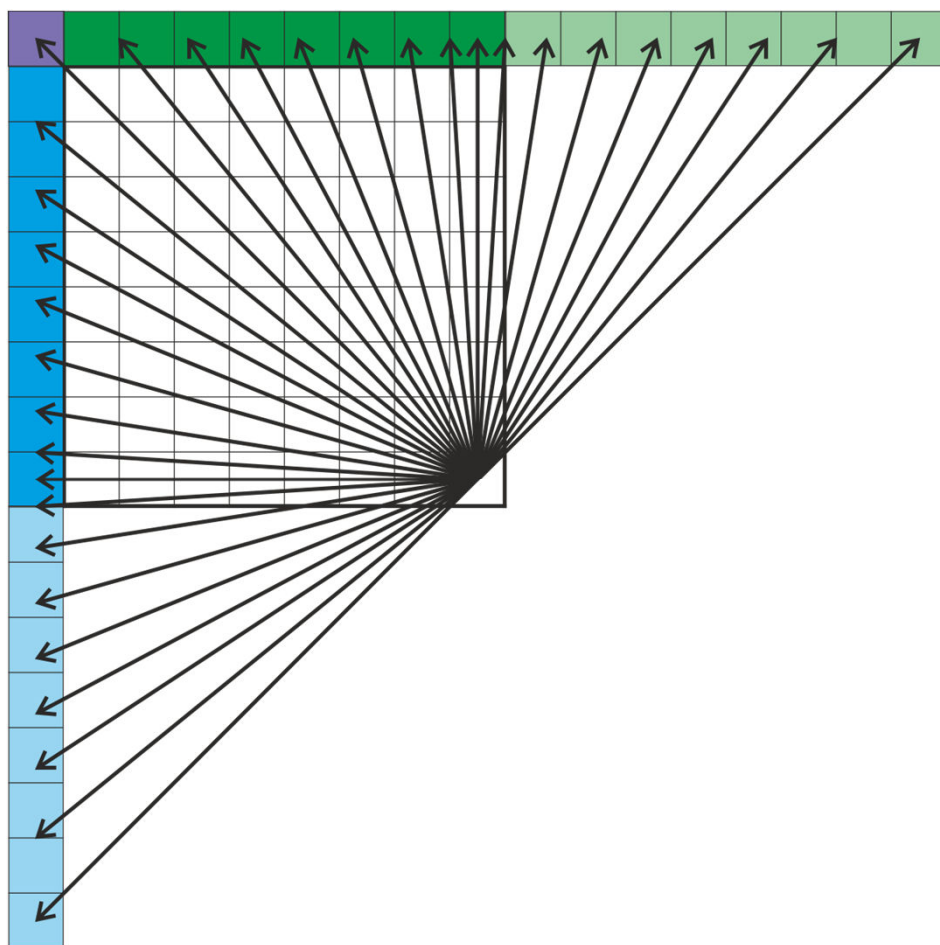


Próbki odniesienia

PUAboveLeft		PUAbove								PUAboveRight							
PULeft																	
PULeftBottom																	

Tryby predykcji wewnątrzobrazowej

- Mode 0: DC (Average value)
- Mode 1: Planar (Bidirectional weighted)
- Mode 2-34: Angular

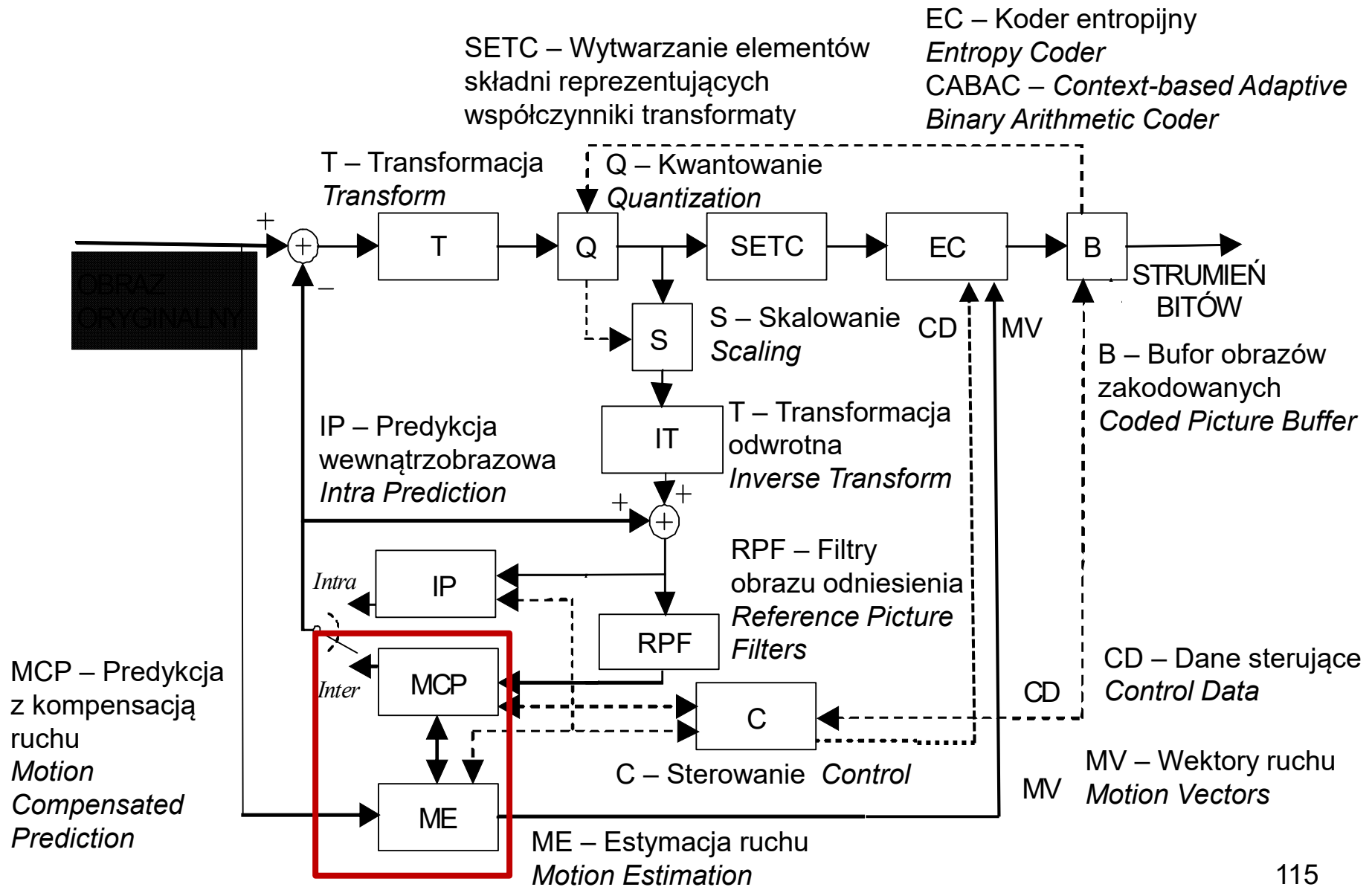


Intra mode signaling

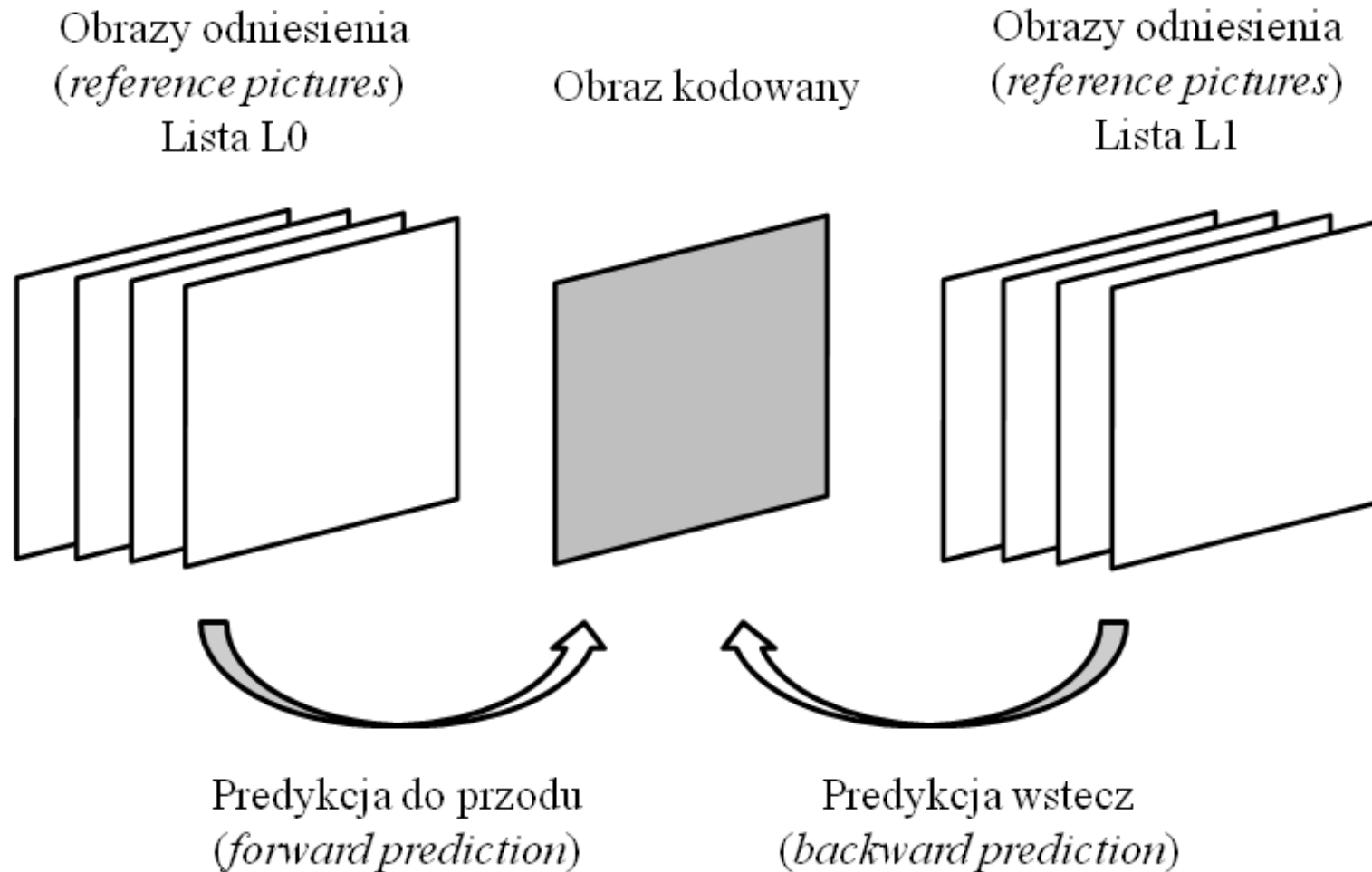
LUMA :

- MPM – Most Probably Modes
Three modes derived from the left and top blocks.
- Index to the mode in MPM
or
- Index to the absolute code of mode .

8.9. Predykcja międzyobrazowa ♦ Predykcja z kompensacją ruchu



Predykcja międzyobrazowa z wieloma obrazami odniesienia



Listy L1 i L0 są ustalane dla całego obrazu.

Predykcja ważona

- Wartości próbek są mnożone przez współczynnik $\alpha \in (0, 1)$.
- - **Jawna predykcja ważona** (*explicit weighted prediction*) w plastrach P i B – wartość współczynnika α jest przesyłana w nagłówku plastra;
- - **Ukryta predykcja ważona** (*implicit weighted prediction*) tylko w plastrach B – wartości współczynników α dla obu predykcji są wyznaczane jako liczby odwrotnie proporcjonalne do długości przedziałów czasu pomiędzy odpowiednim obrazem odniesienia a obrazem kodowanym.

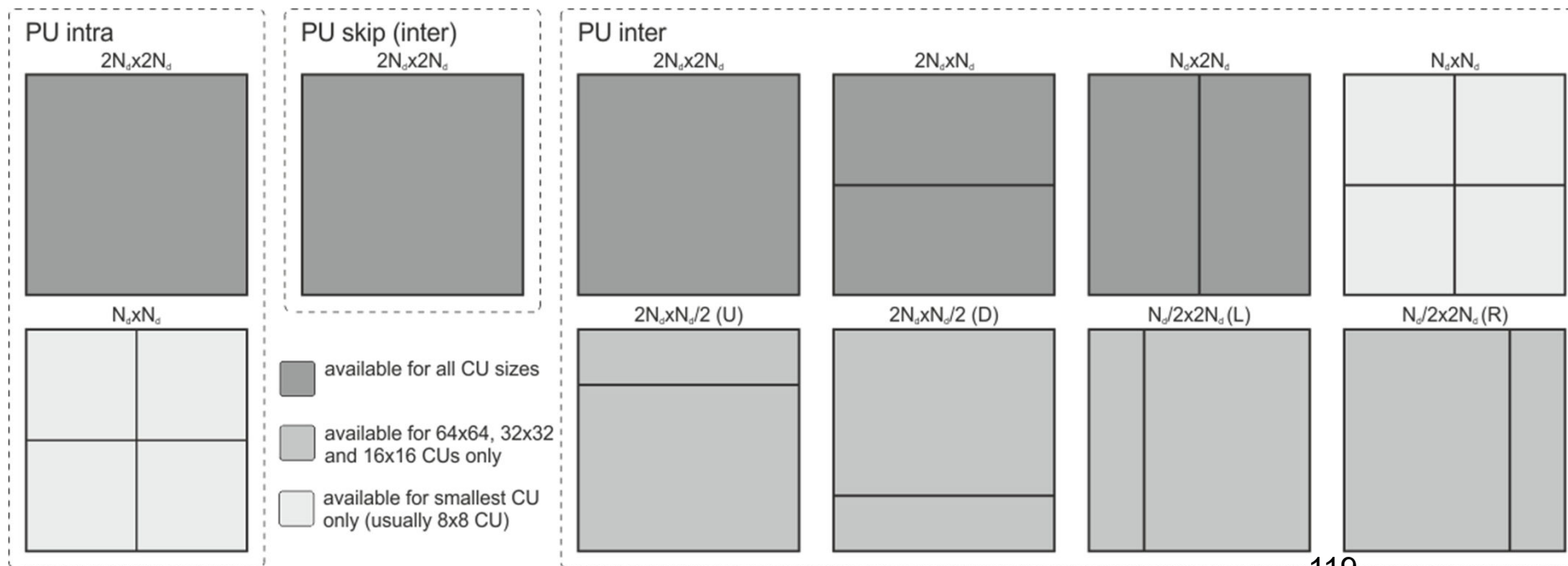
Predykcja międzyobrazowa ♦ Predykcja z kompensacją ruchu

- Wektory ruchu z dokładnością $\frac{1}{4}$ okresu próbkowania.
- **Interpolacja obrazu odniesienia** (w położeniach $\frac{1}{2}$, $\frac{1}{4}$ i $\frac{3}{4}$ okresu próbkowania).
- Próbki obrazu (A)
- Próbki interpolowane (b, h, j)
- (a, c, d, e, g, i, k, n, p, q, r)

$A_{-1,-1}$				$A_{0,-1}$	$a_{0,-1}$	$b_{0,-1}$	$c_{0,-1}$	$A_{1,-1}$				$A_{2,-1}$
$A_{-1,0}$				$A_{0,0}$	$a_{0,0}$	$b_{0,0}$	$c_{0,0}$	$A_{1,0}$				$A_{2,0}$
$d_{-1,0}$				$d_{0,0}$	$e_{0,0}$	$f_{0,0}$	$g_{0,0}$	$d_{1,0}$				$d_{2,0}$
$h_{-1,0}$				$h_{0,0}$	$i_{0,0}$	$j_{0,0}$	$k_{0,0}$	$h_{1,0}$				$h_{2,0}$
$n_{-1,0}$				$n_{0,0}$	$p_{0,0}$	$q_{0,0}$	$r_{0,0}$	$n_{1,0}$				$n_{2,0}$
$A_{-1,1}$				$A_{0,1}$	$a_{0,1}$	$b_{0,1}$	$c_{0,1}$	$A_{1,1}$				$A_{2,1}$
$A_{-1,2}$				$A_{0,2}$	$a_{0,2}$	$b_{0,2}$	$c_{0,2}$	$A_{1,2}$				$A_{2,2}$

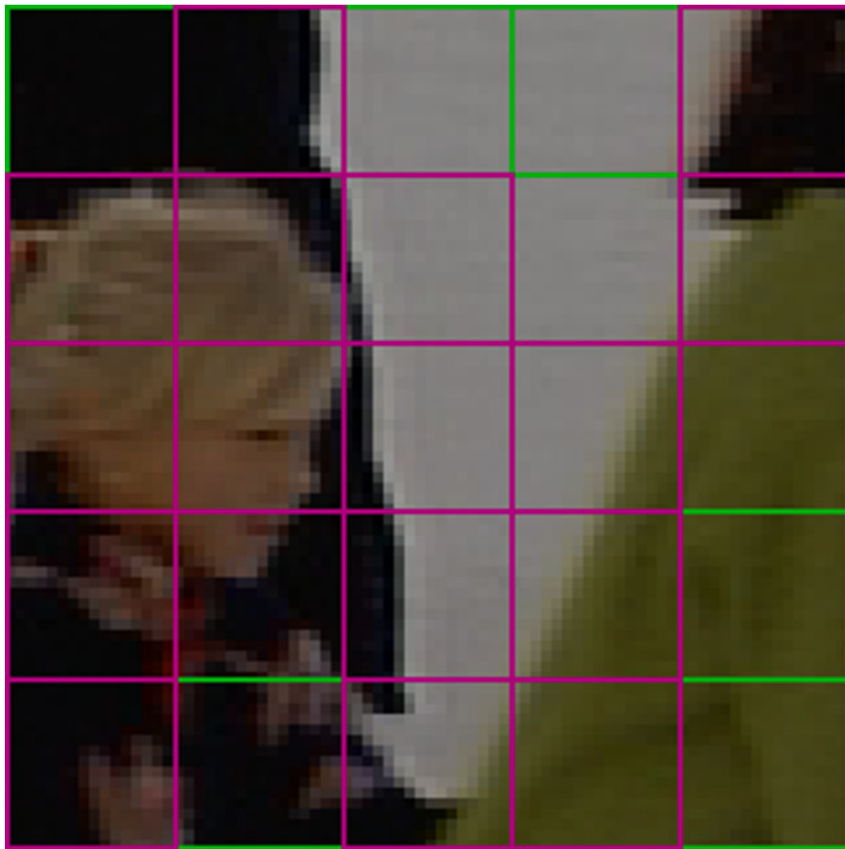
Jednostki predykcji

- CU zawiera jedną lub więcej jednostek predykcji PUs
- Wyjątki:
 - Jednostki predykcji rozmiaru 4x4 są zakazane.
 - Jednostki predykcji rozmiaru 8x4 i 4x8 mogą służyć tylko predykcji jednokierunkowej.



Predykcja blokami o zmiennych rozmiarach

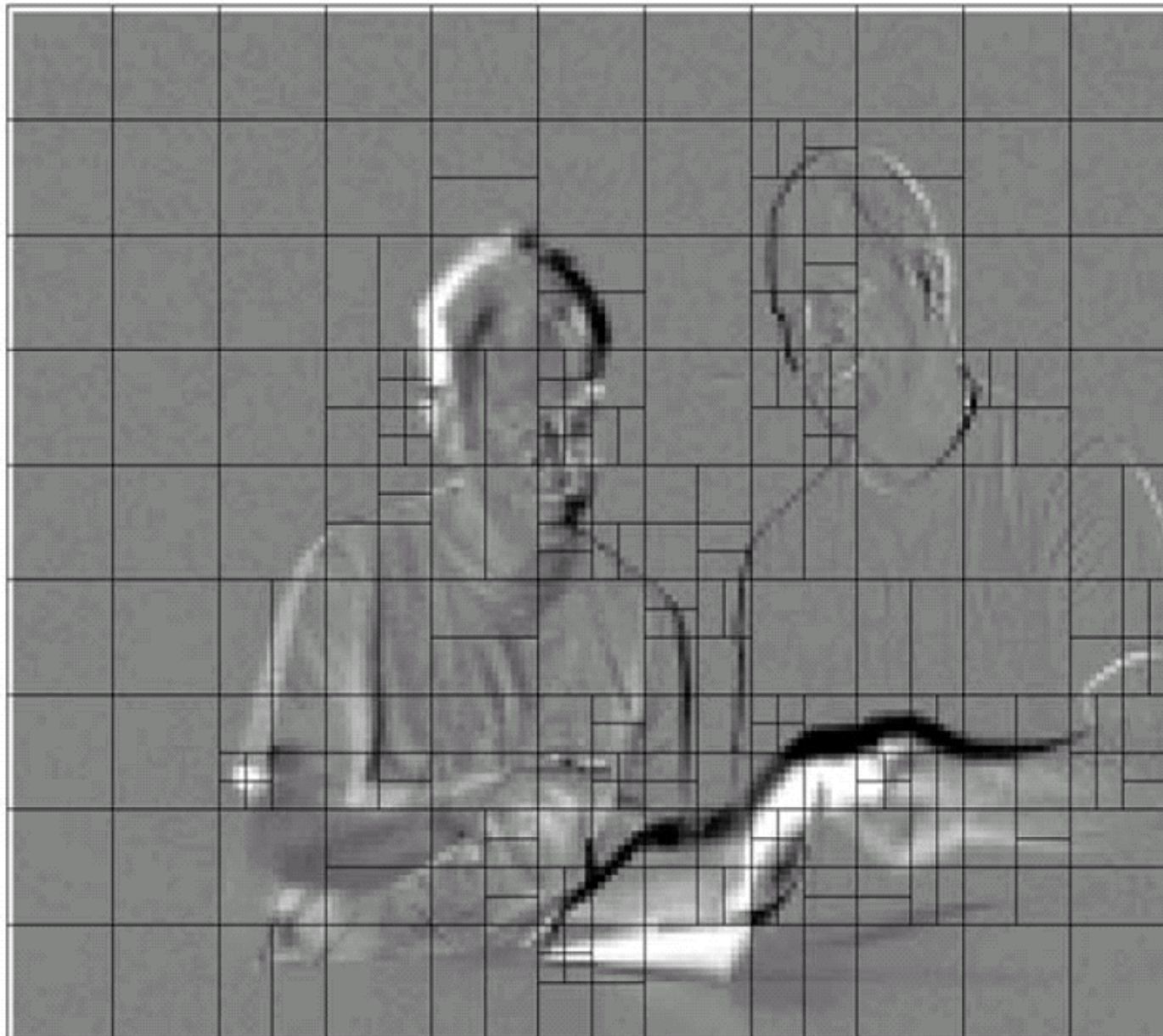
Stały rozmiar bloku



Jak w HEVC

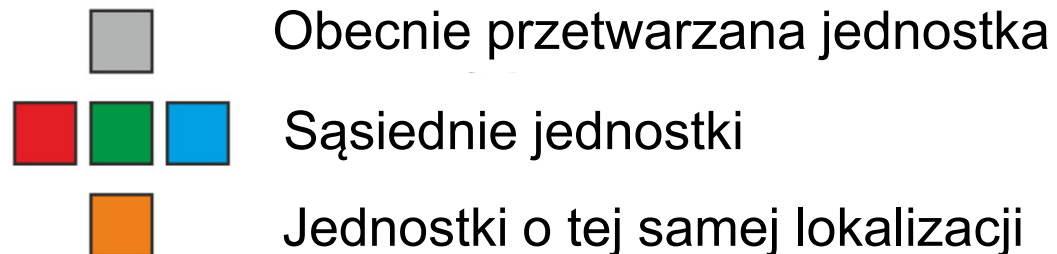
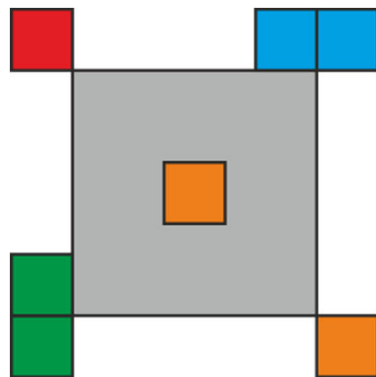


Predykcja blokami o zmiennych rozmiarach



Predykcja wektorów ruchu

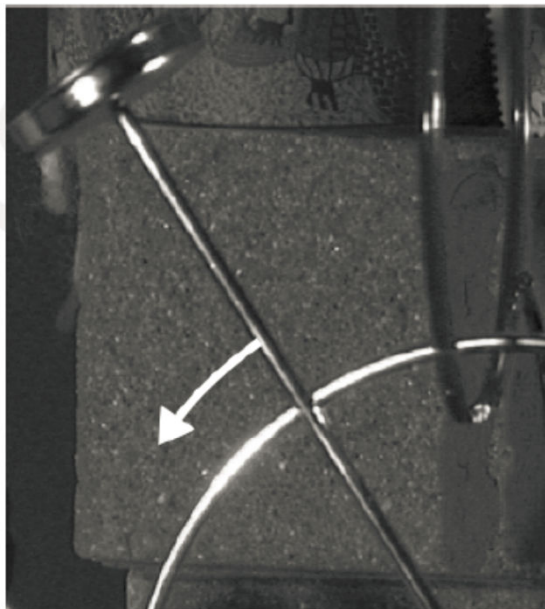
- Konkurencja wektorów ruchu – Motion vector competition
- Scalanie wektorów ruchu – Motion data merging



Ewentualnie przesyłane są błędy predykcji składowych wektorów ruchu.

Tryb scalania (*Merge mode*)

Scalanie jednostek predykcji PU



(a)



(b)



(c)

Poruszający się obiekt

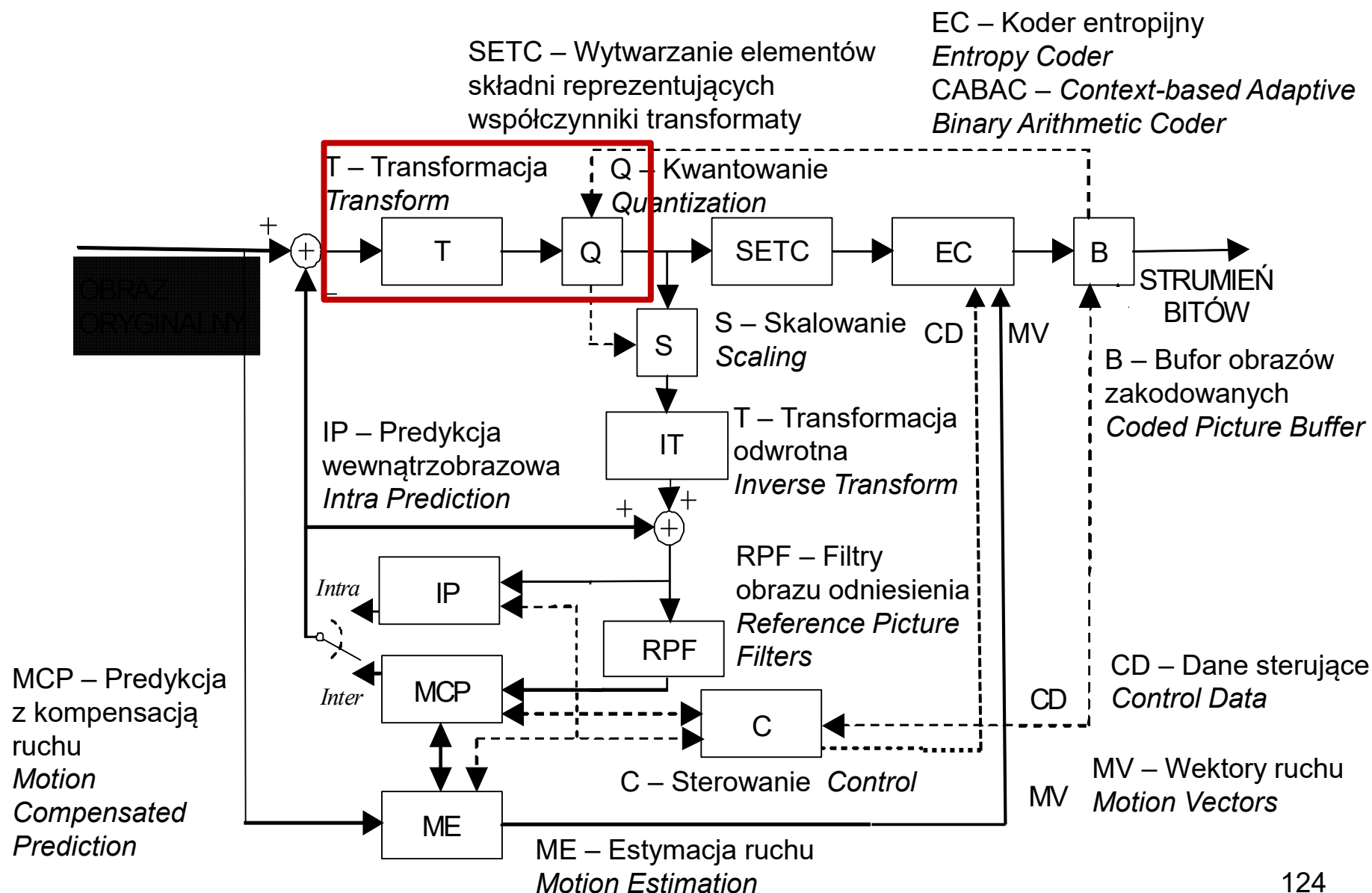
Bez trybu scalania

Z trybem scalania

(znacznie więcej parametrów ruchu)

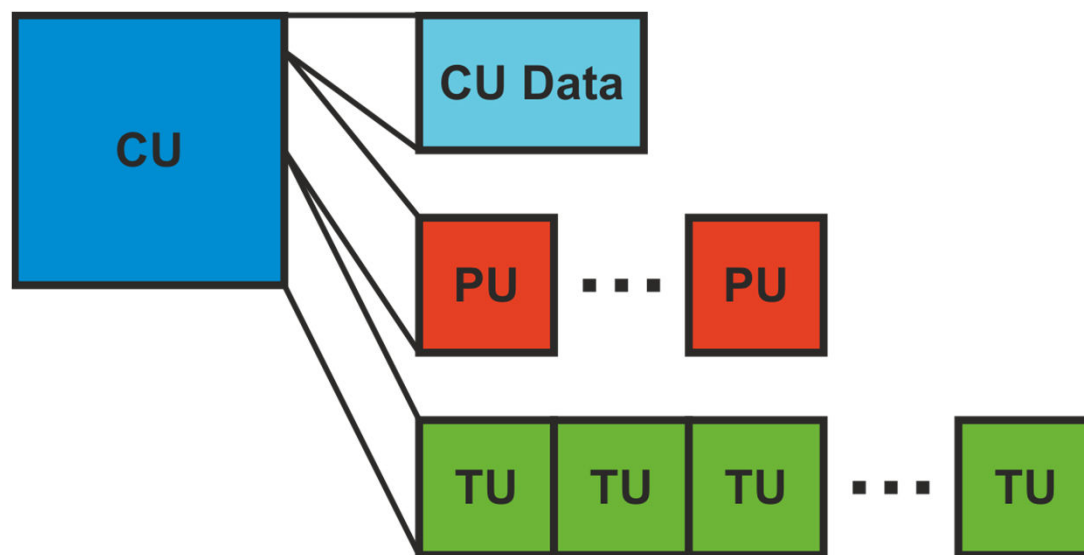
Tryb scalania przestrzennego i czasowego

8.10. Kodowanie błędu predykcji



Jednostki transformacji

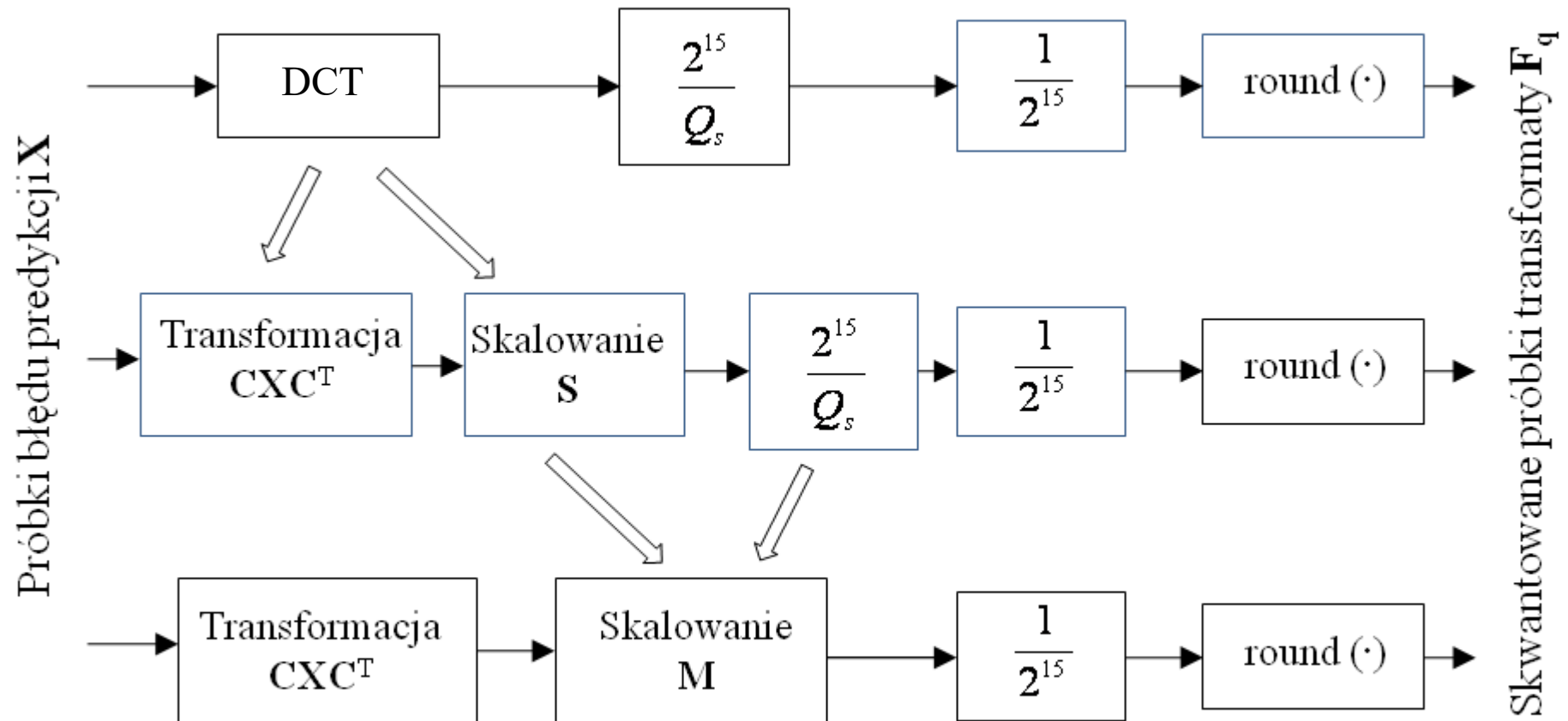
- Podział jednostek kodowania (CU) na jednostki transformacji (*transform units* – TU).
- Jednostki transformacji są kwadratowe o rozmiarze 4 x 4, 8 x 8, 16 x 16 lub 32 x 32.



Kodowanie błędu predykcji: transformacje

- Transformacje:
 - DCT (Discrete Cosine Transform) – 4x4, 8x8, 16x16 and 32x32.
 - DST (Discrete Sine Transform) – Intra 4x4 tylko.
 - TS (Transform Skip) mode – 4x4 only (**no transformation !!!**).
- Kwantowanie:
 - Krok stratny !
 - QP (Parametr kwantyzacji - Quantization parameter): 0-51;
 - 0 – smallest quality loss, highest bitrate,
 - 51 – highest quality loss, smallest bitrate.

Transformacja kosinusowa i kwantowanie



Bez tablicy skalowania:

$$\mathbf{M} = \frac{2^{15}}{Q_s} \cdot \mathbf{S}$$

Macierz $\mathbf{M}_H$ zastępuje
 $\mathbf{M}$

Z tablicą skalowania:

w schemacie blokowym

$$M_H(k_1, k_2) = \frac{16 \cdot M(k_1, k_2)}{Q_T(k_1, k_2)}$$

Przykładowa tablica Q_T skalująca dla bloków 4×4 kodowanych wewnątrzobrazowo

6	13	20	28
13	20	28	32
20	28	32	37
28	32	37	42

Tablica M_H

Przykładowa tablica Q_T skalująca dla bloków 4×4 kodowanych międzyobrazowo

10	14	20	28
14	20	24	27
20	24	27	30
28	27	30	34

Tablica M_H

Kwantowanie

$$F_q(i, j) = \text{round} \left(\frac{F(i, j)}{Q_{step}} \right)$$

<i>QP</i>	0	1	2	3	4	5	6	7	8	9	10	11	12	...
<i>QStep</i>	0.625	0.6875	0.8125	0.875	1	1.125	1.25	1.375	1.625	1.75	2	2.25	2.5	...
<i>QP</i>	...	18	...	24	...	30	...	36	...	42	...	48	...	51
<i>QStep</i>		5		10		20		40		80		160		224



Oznaczane dalej jako: Q_s

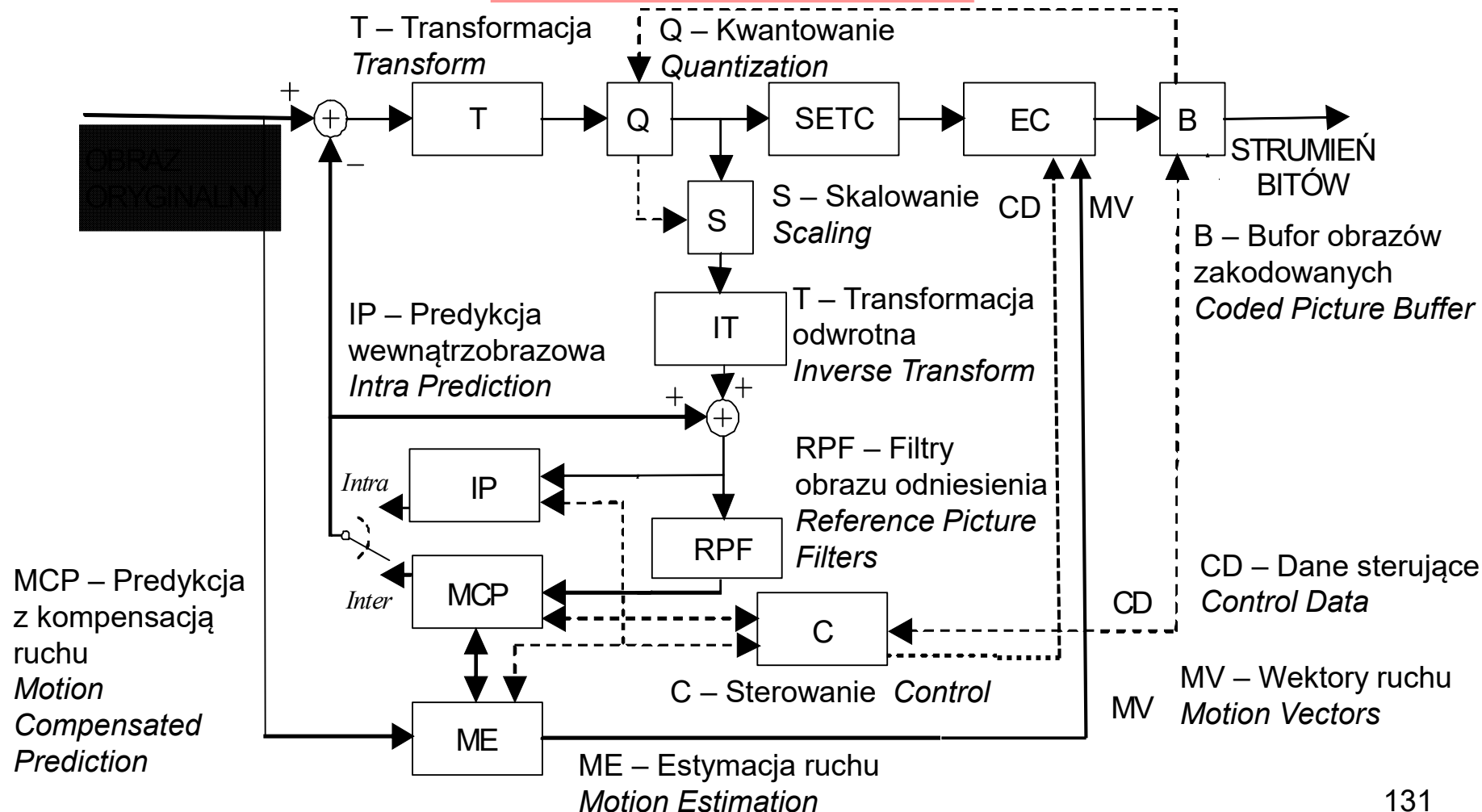
Kwantowanie

- Współczynnik QP jest głównym parametrem pozwalającym na regulowanie prędkości bitowej.
- Zmiana współczynnika QP zmienia wielkości kroku kwantowania dla wszystkich częstotliwości przestrzennych.
- Krok kwantowania może być stały dla wszystkich częstotliwości przestrzennych, a może być zróżnicowany i zdefiniowany dla różnych częstotliwości przestrzennych w tablicy skalującej $\mathbf{M}_H$.

Przesyłanie skwantowanych współczynników transformaty do kodera entropijnego

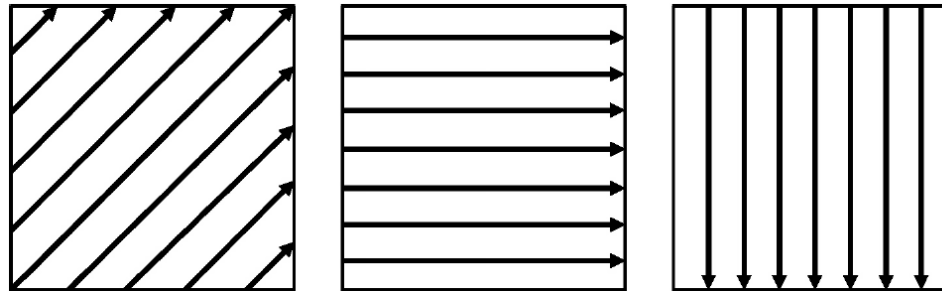
SETC – Wytwarzanie elementów
składni reprezentujących
współczynniki transformaty

EC – Koder entropijny
Entropy Coder
CABAC – *Context-based Adaptive
Binary Arithmetic Coder*



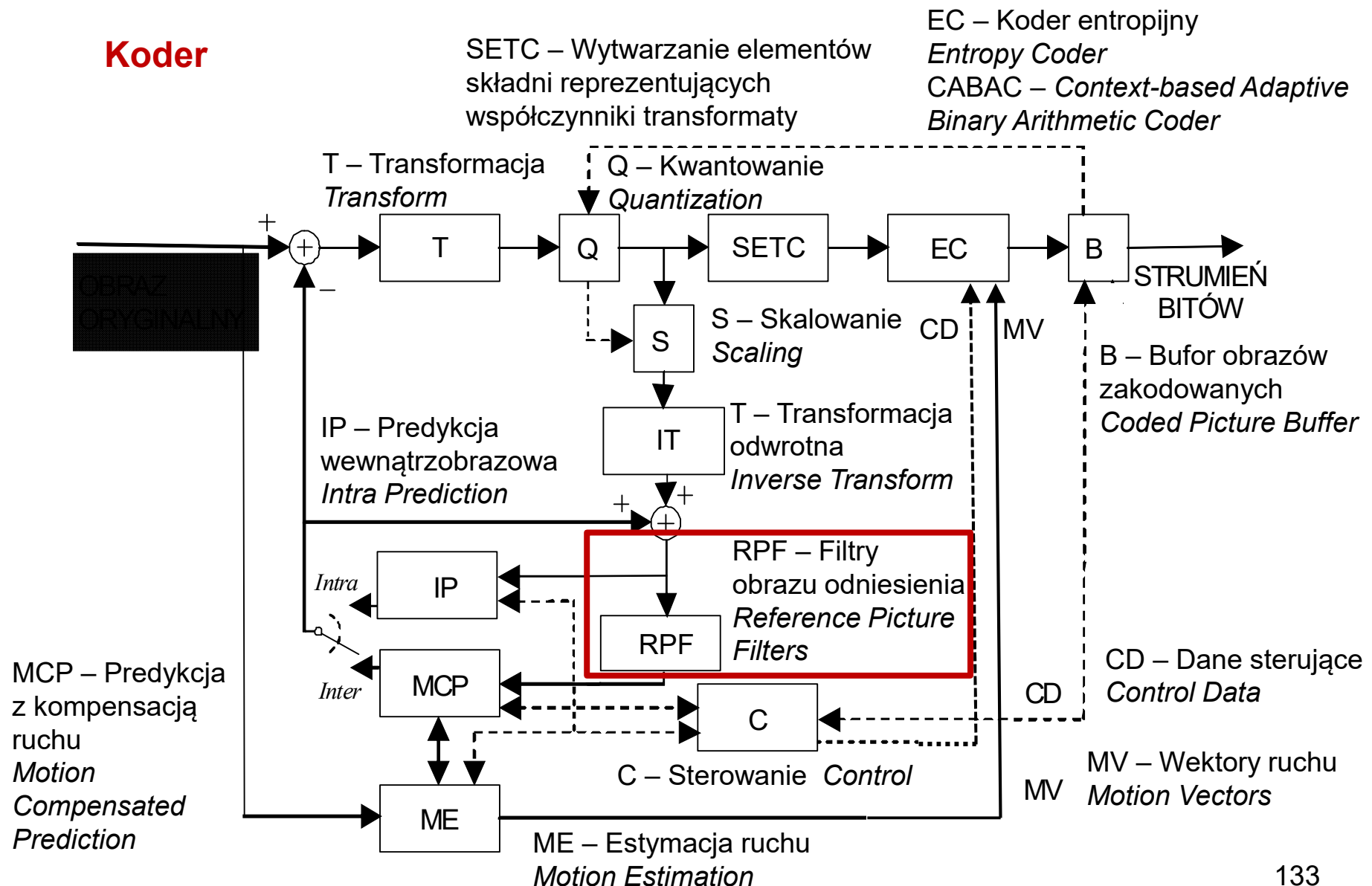
Kolejne przesyłanie skwantowanych współczynników transformaty do kodera entropijnego

- 3 typy porządkowania: diagonalny, poziomy i pionowy.

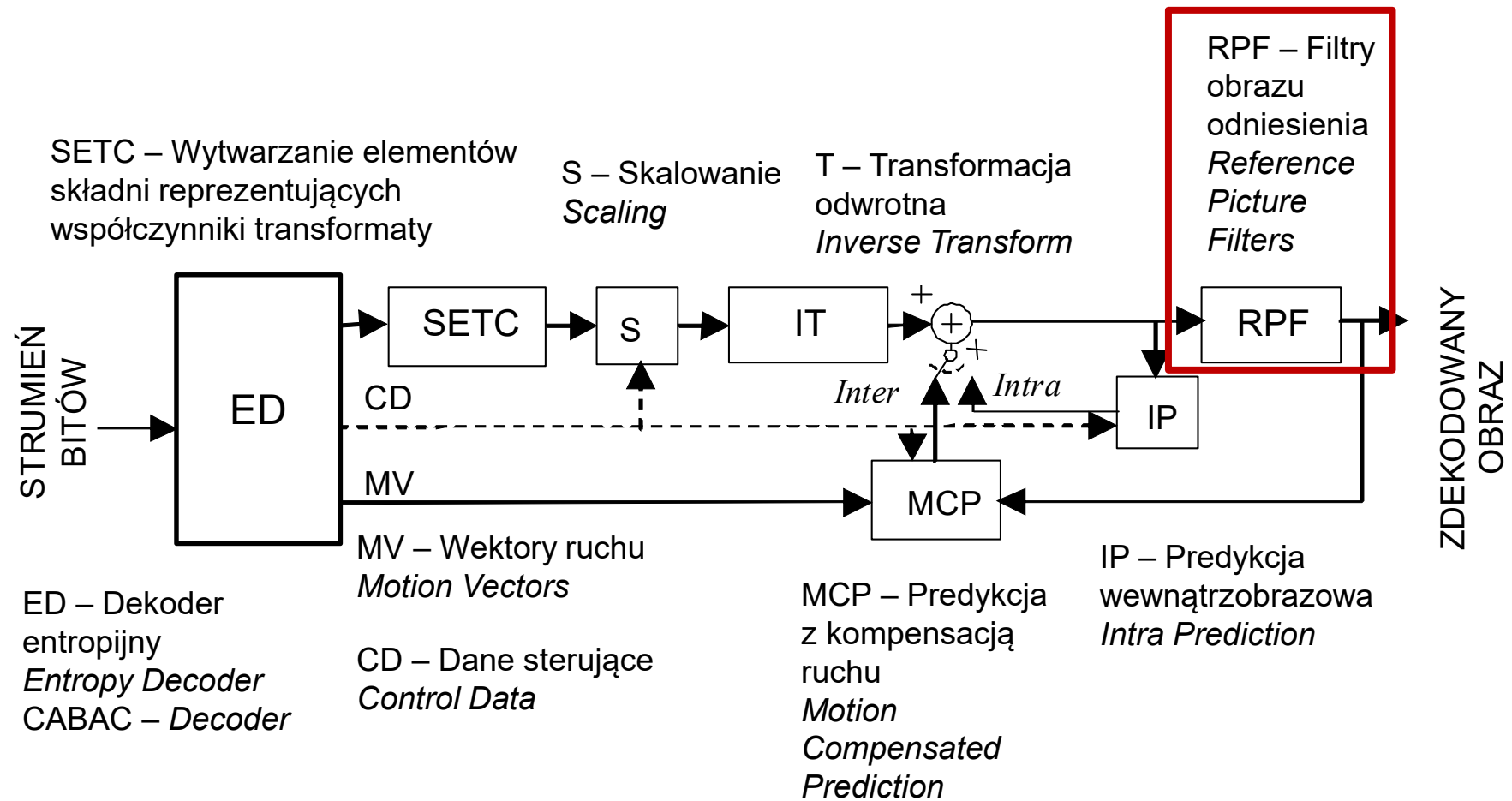


8.11. Filtracja obrazów odniesienia (*in-loop filters*)

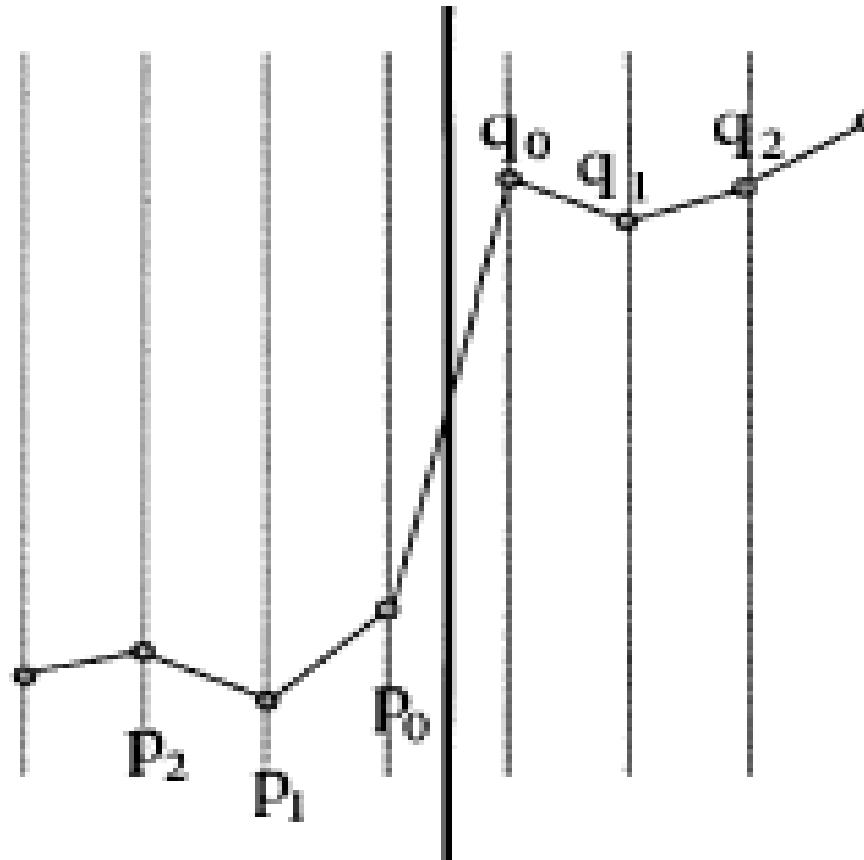
Koder



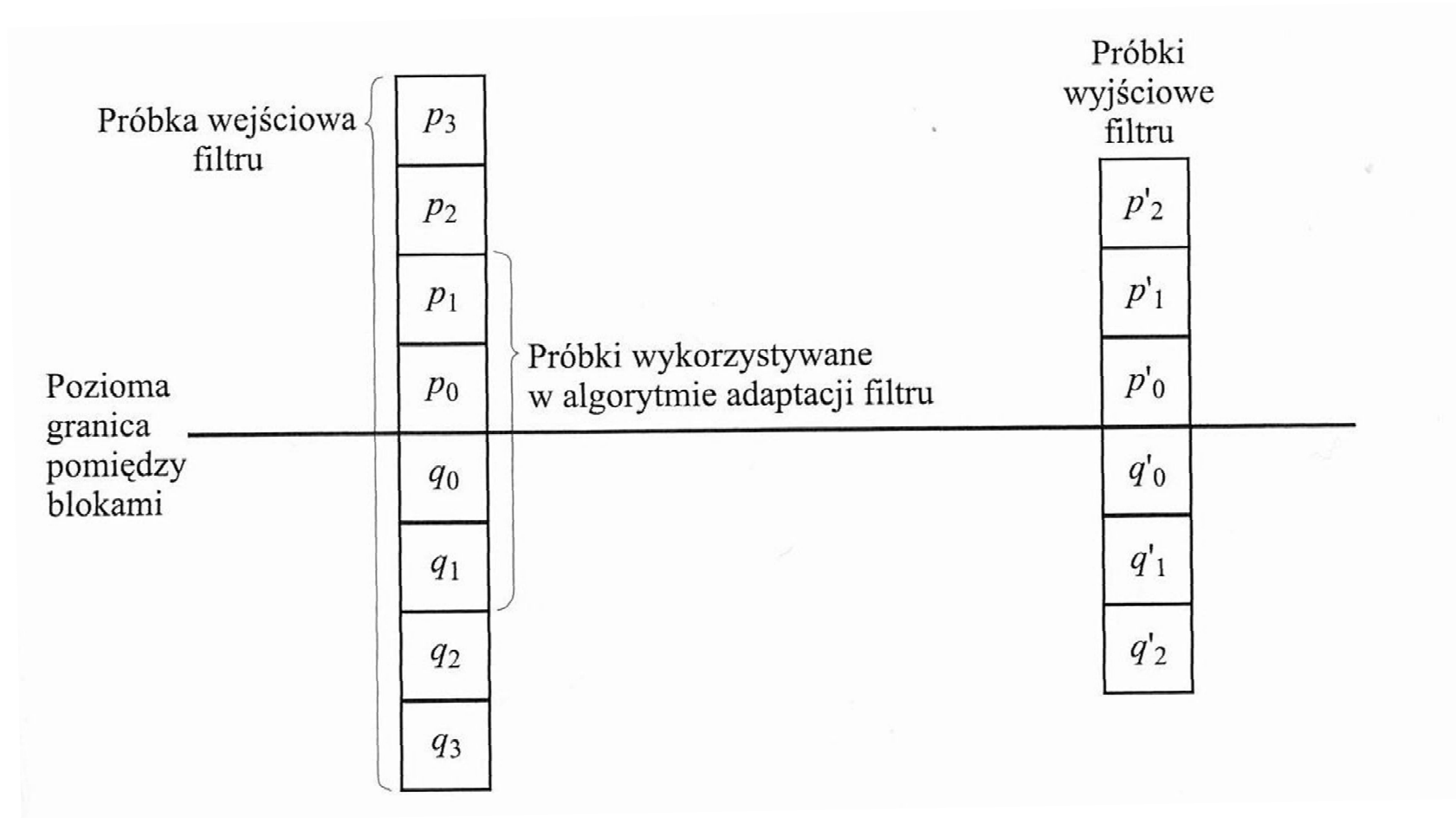
Dekoder HEVC



Filtr redukujący efekt blokowy w obrazie odniesienia

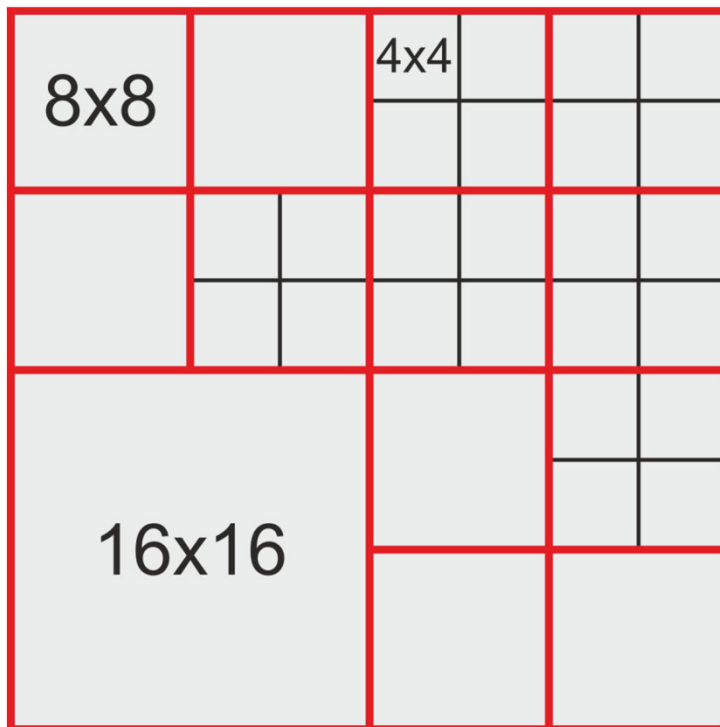


Filtr redukujący efekty blokowe

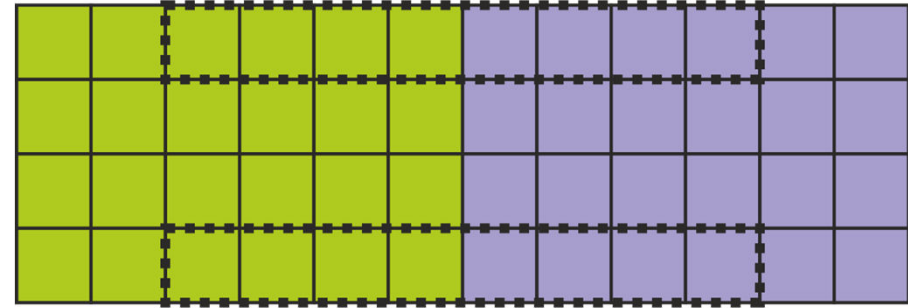


Filtr redukujący efekty blokowe

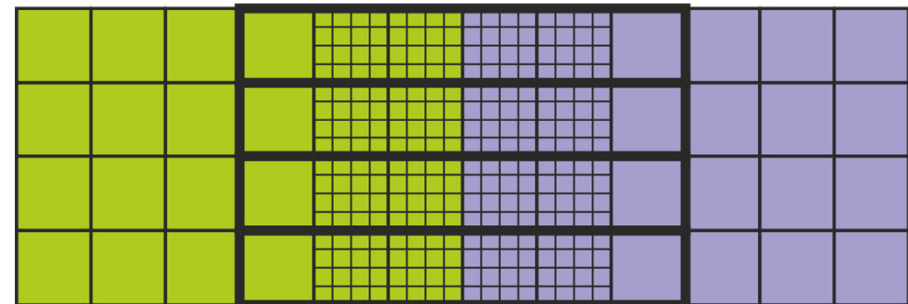
- HEVC stosuje filtr deblokujący na krawędziach siatki 8×8 .
- Możliwe zrównoleglenie.



Decyzja o filtracji (tak/nie)



Słaby filtr



Silny filtr

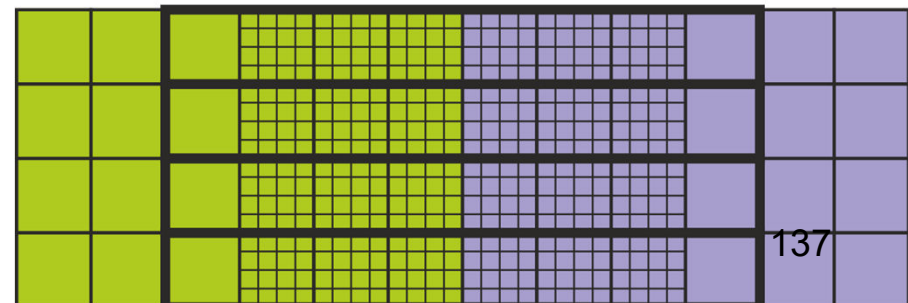




Figure 6.33 Reconstructed, $QP = 36$ (no filter)



Figure 6.34 Reconstructed, $QP = 36$ (with filter)



Figure 6.35 Reconstructed, $QP = 32$ (no filter)



Figure 6.36 Reconstructed, $QP = 32$ (with filter)

Wycinek z oryginalnej sekwencji, 704 x 576, 25 obraz/s



Sekwencja zdekodowana, filtr włączony



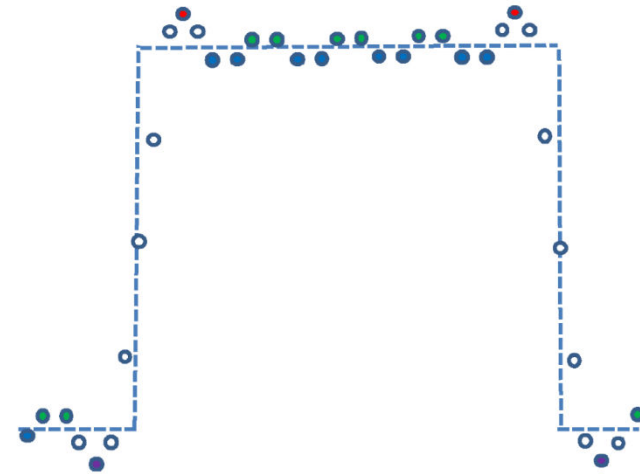
Sekwencja zdekodowana (1 Mb/s), filtr wyłączony



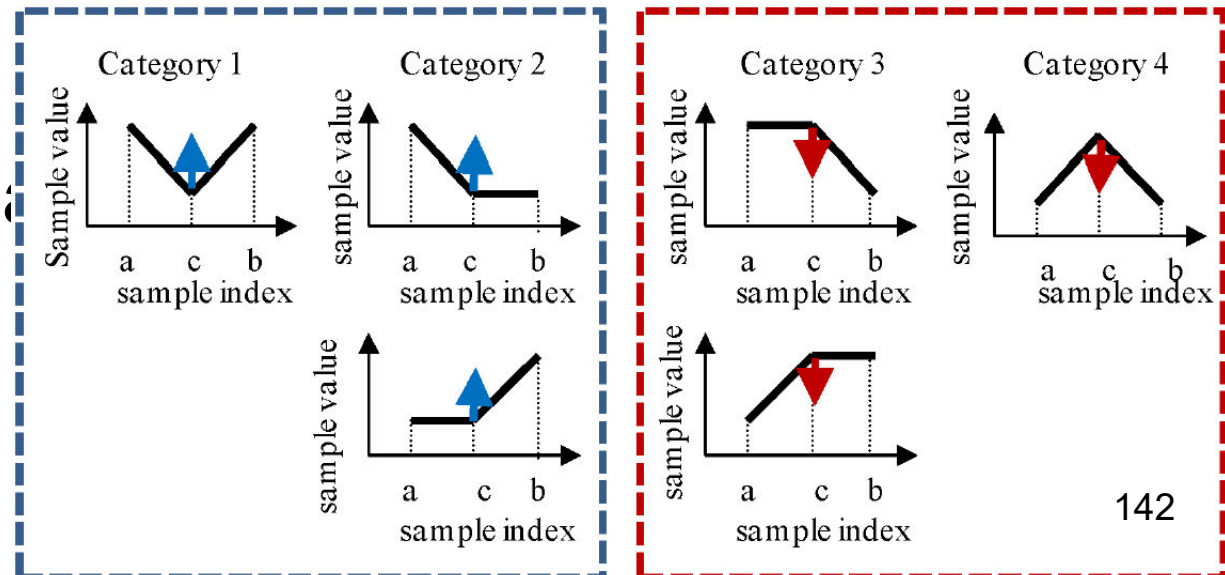
Zmiana wartości próbek SAO (Sample Adaptive Offset)

- Redukuje dzwonienia w obrazie odniesienia.

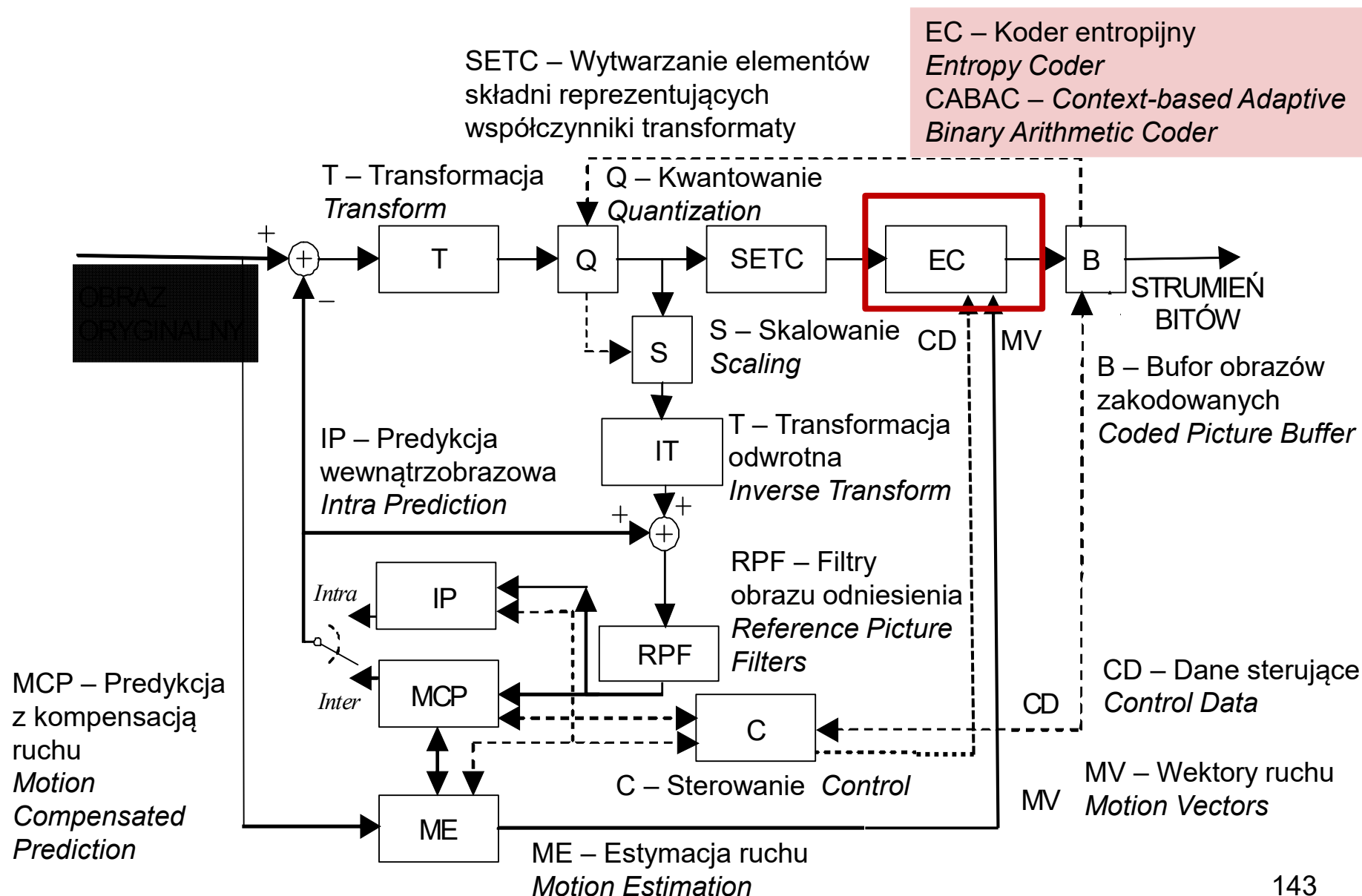
- Dzwonienia w obrazie odniesienia



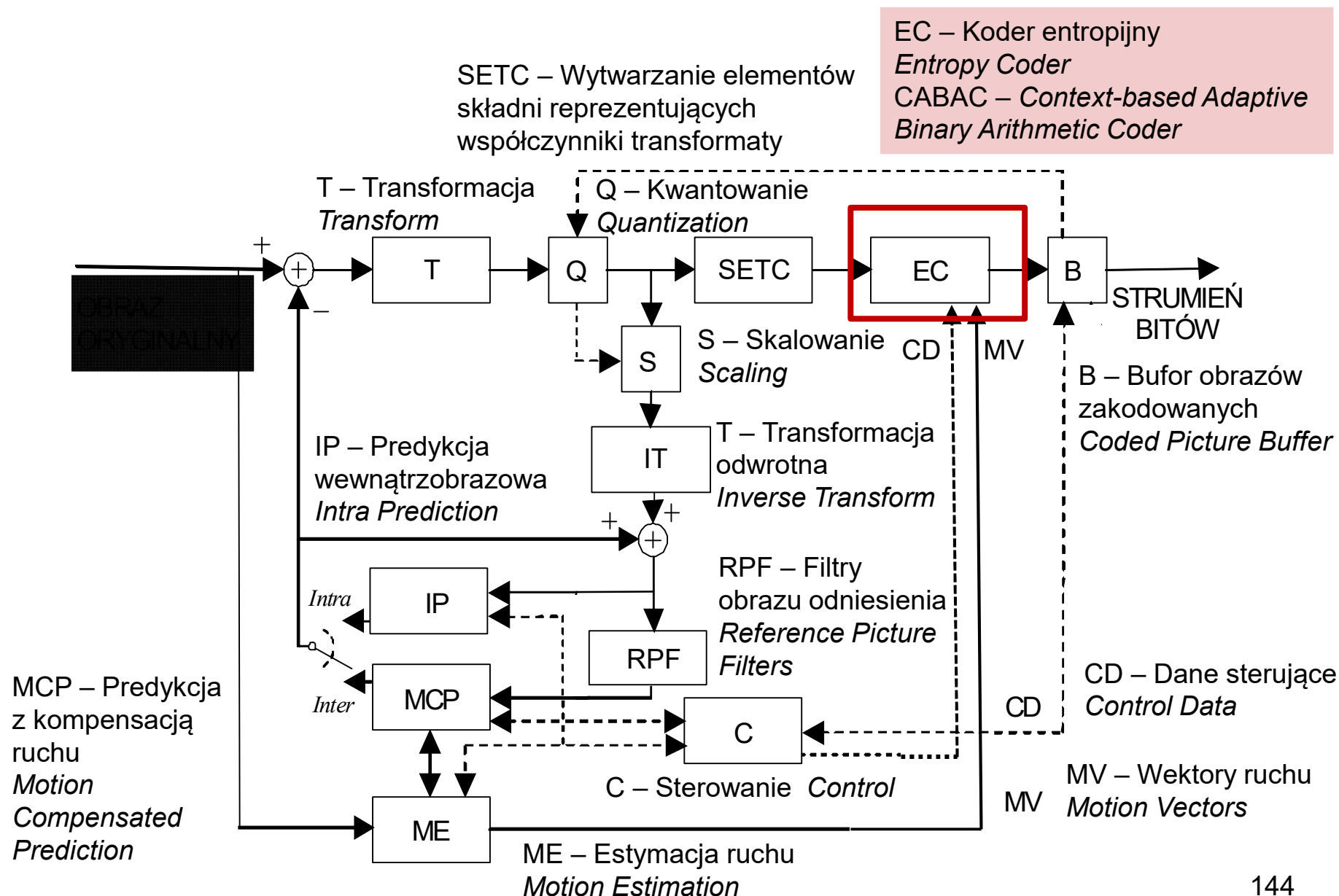
- Ich redukcja:



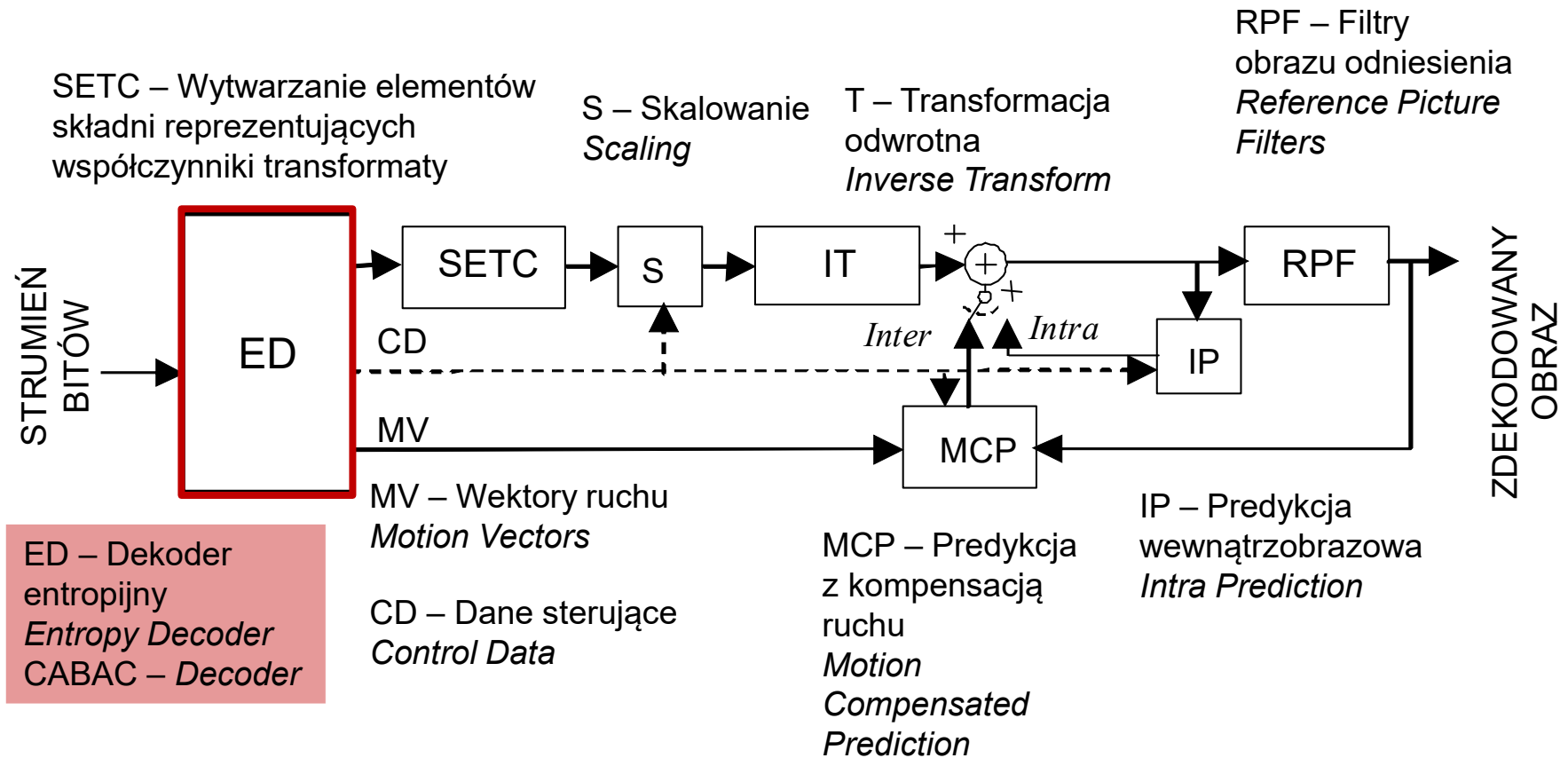
8.12. Kodowanie entropijne



8.12. Kodowanie entropijne

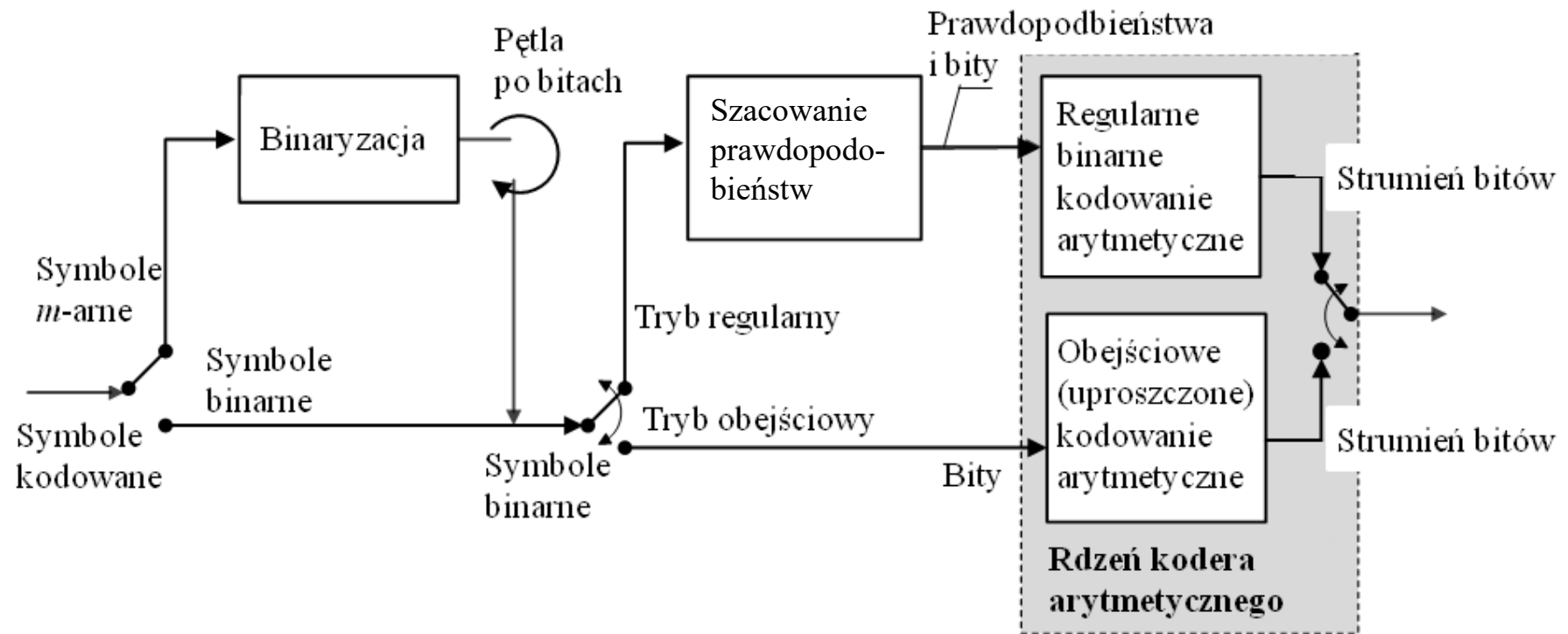


Dekodowanie entropijne



Kontekstowe adaptacyjne binarne kodowanie arytmetyczne CABAC

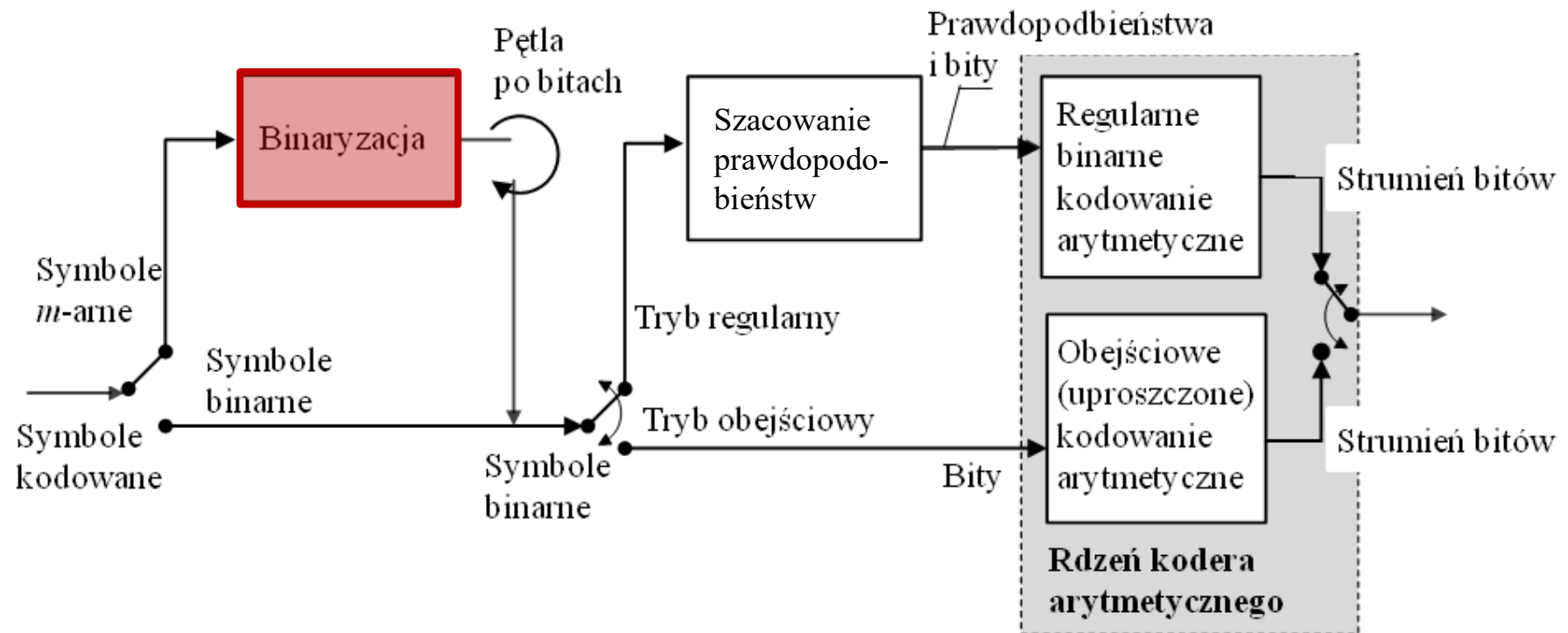
Context-based Adaptive Binary Arithmetic Coding



CABAC

Context-based Adaptive Binary Arithmetic Coding

Kody uniwersalne



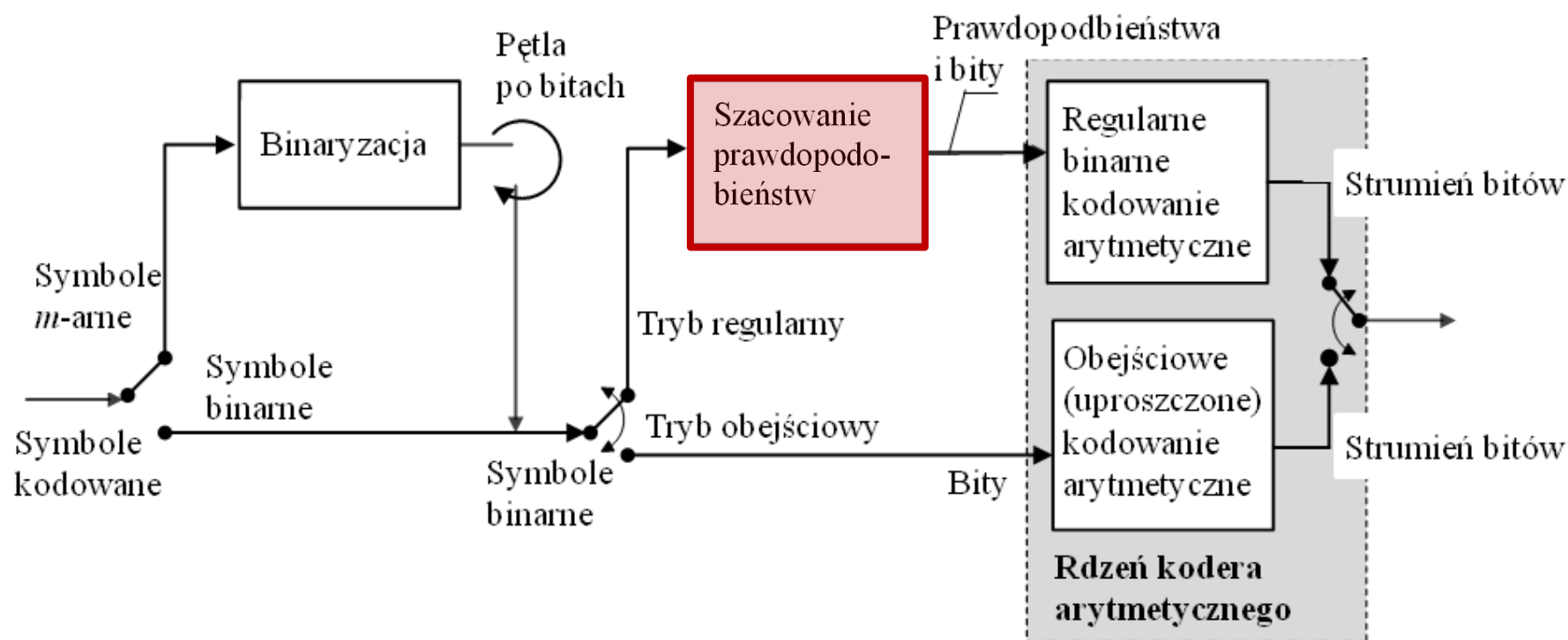
Binaryzacja

Wykładniczy kod Golomba

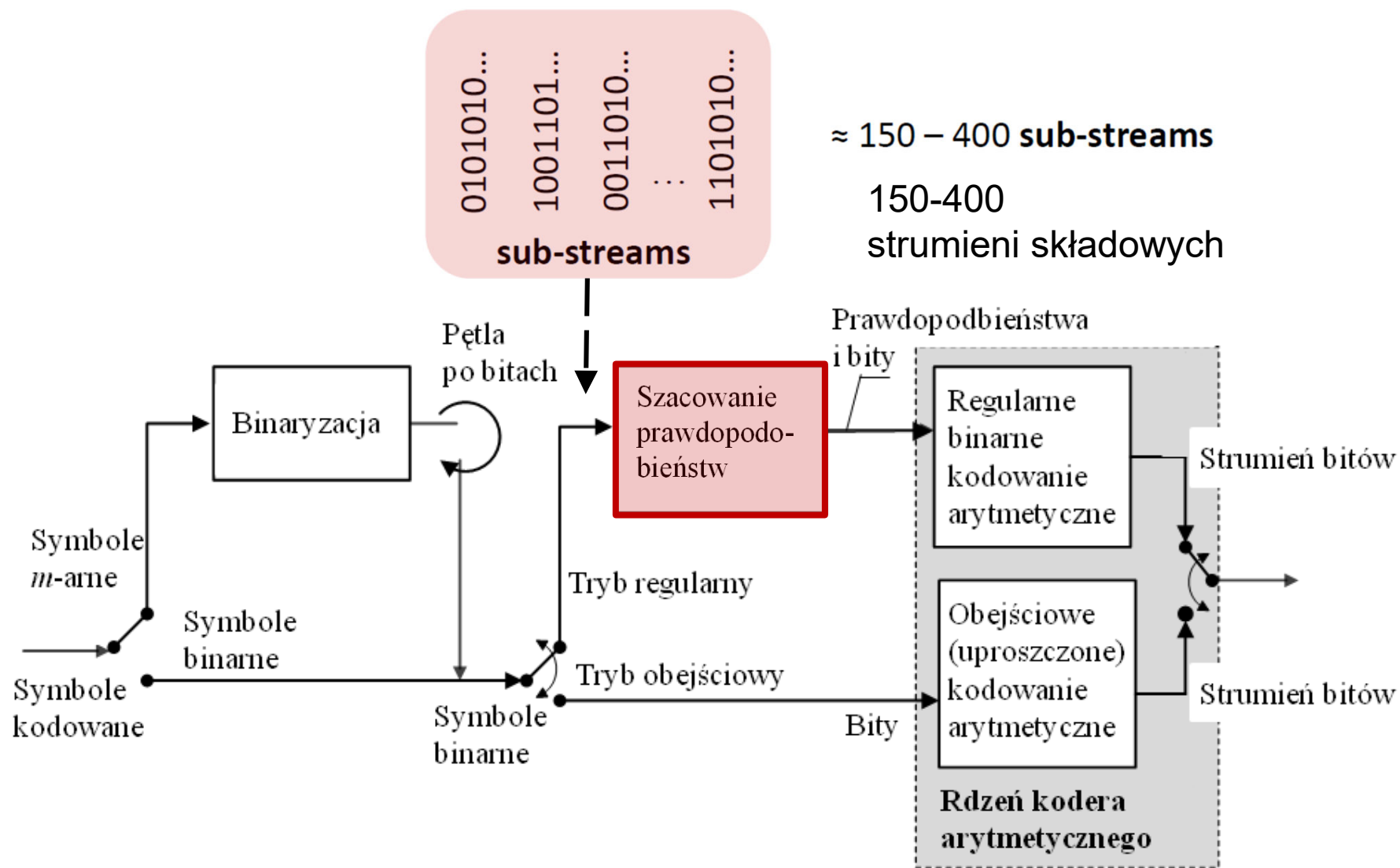
Słowo kodowe	n
1	0
0 1 x_0	1-2
0 0 1 $x_1 x_0$	3-6
0 0 0 1 $x_2 x_1 x_0$	7-14
0 0 0 0 1 $x_3 x_2 x_1 x_0$	15-30
0 0 0 0 0 1 $x_4 x_3 x_2 x_1 x_0$	31-62
...	...

Szacowanie prawdopodobieństw

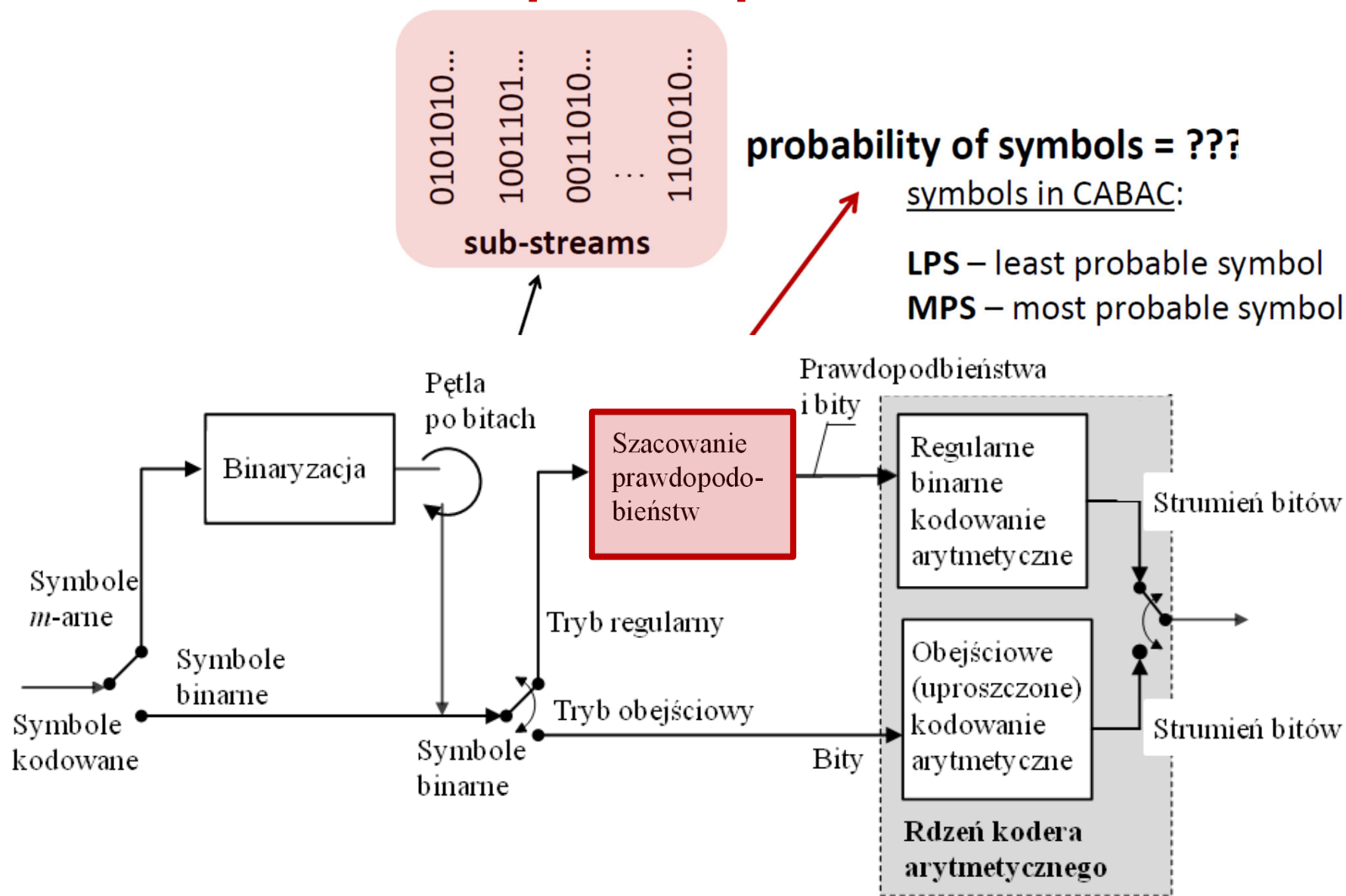
- Odbywa się w plastrach.
- Na początku plastra wszystkie wartości prawdopodobieństw są zakładane zgodnie z wartościami początkowymi określonymi w normie (takimi samymi w koderze i w dekodерze).



Szacowanie prawdopodobieństw

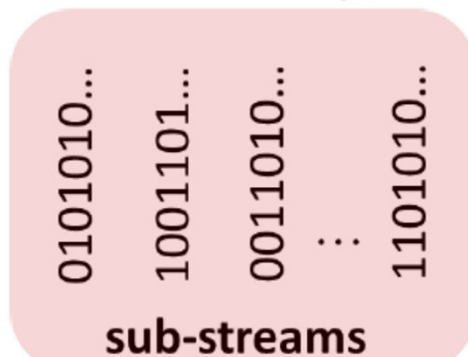


Szacowanie prawdopodobieństw



Szacowanie prawdopodobieństw

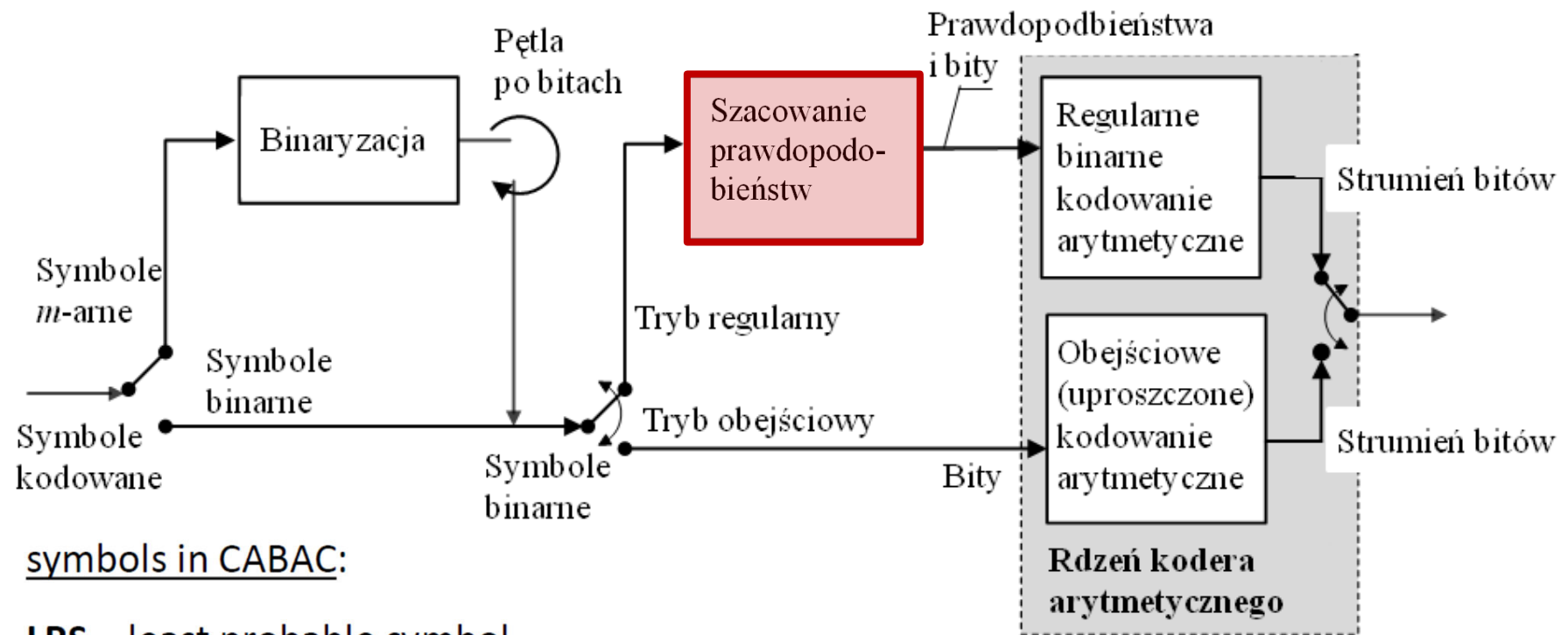
probability of symbols:



$$p_{new}(LPS) = \alpha \cdot y(bin) + (1 - \alpha) \cdot p_{old}(LPS)$$

$p_{new}(LPS)$ is the estimated probability of LPS symbol,

$p_{old}(LPS)$ is the probability of LPS symbol calculated earlier,

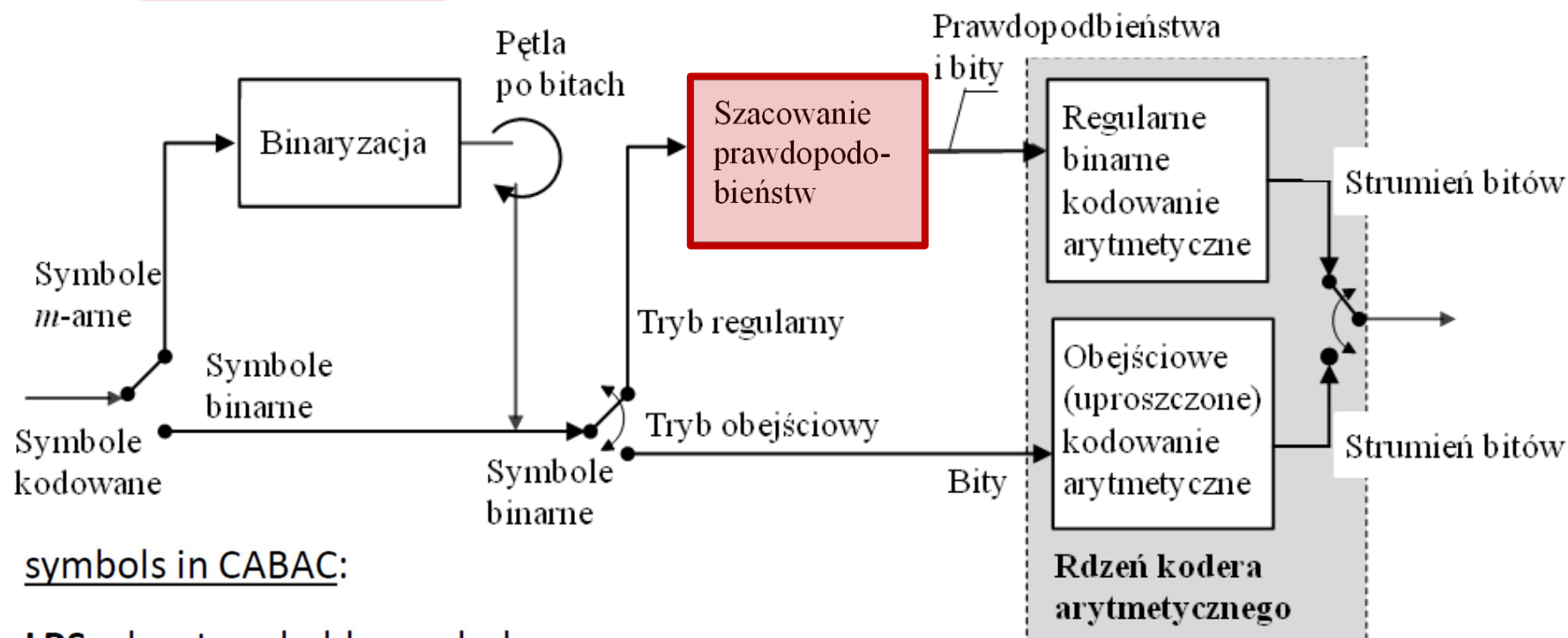
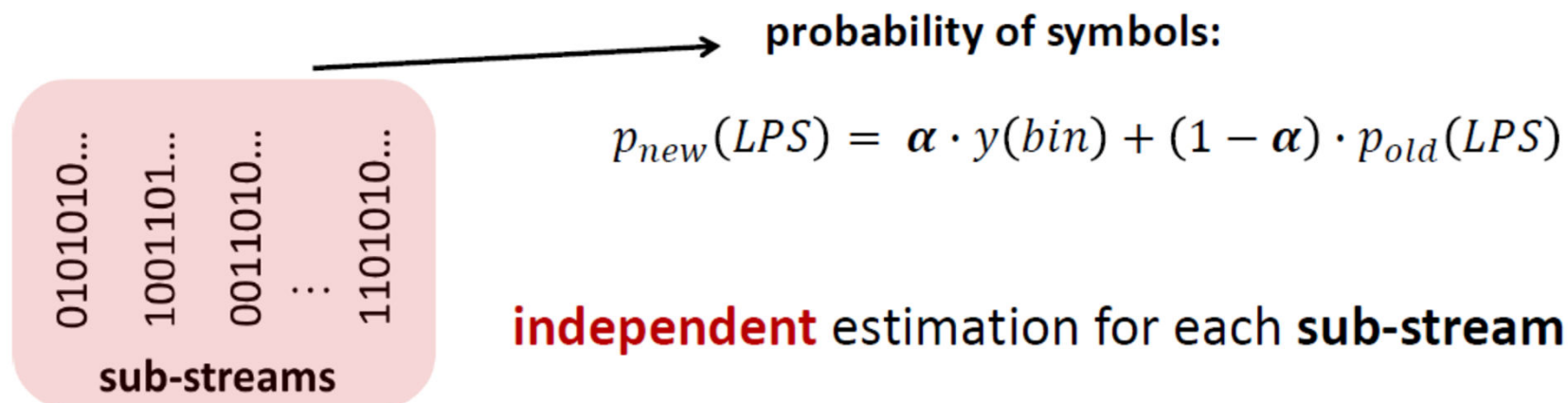


symbols in CABAC:

LPS – least probable symbol

MPS – most probable symbol

Szacowanie prawdopodobieństw

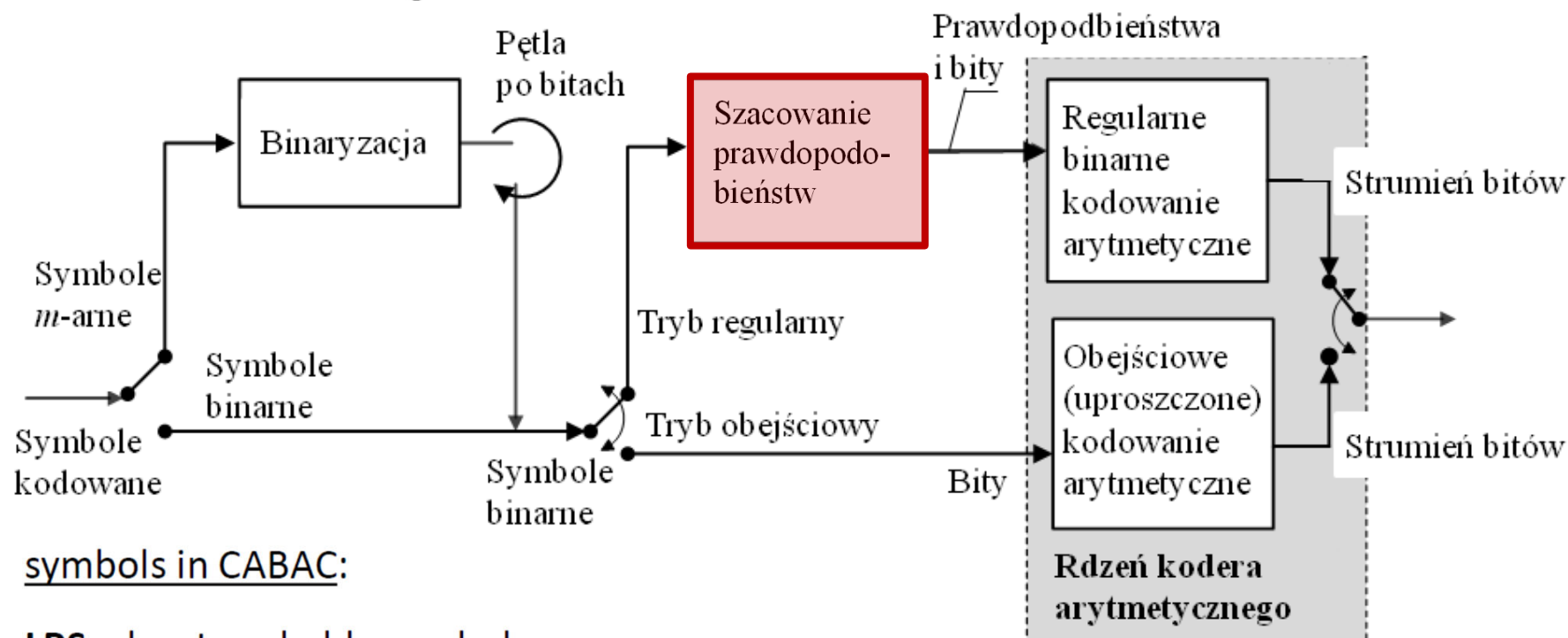
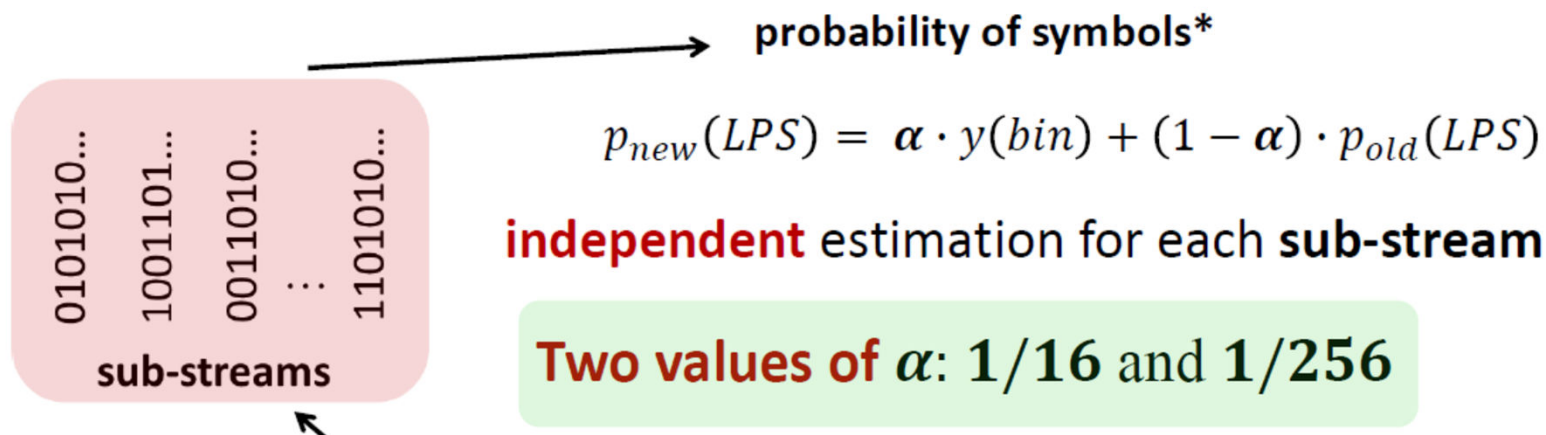


symbols in CABAC:

LPS – least probable symbol

MPS – most probable symbol

Szacowanie prawdopodobieństw

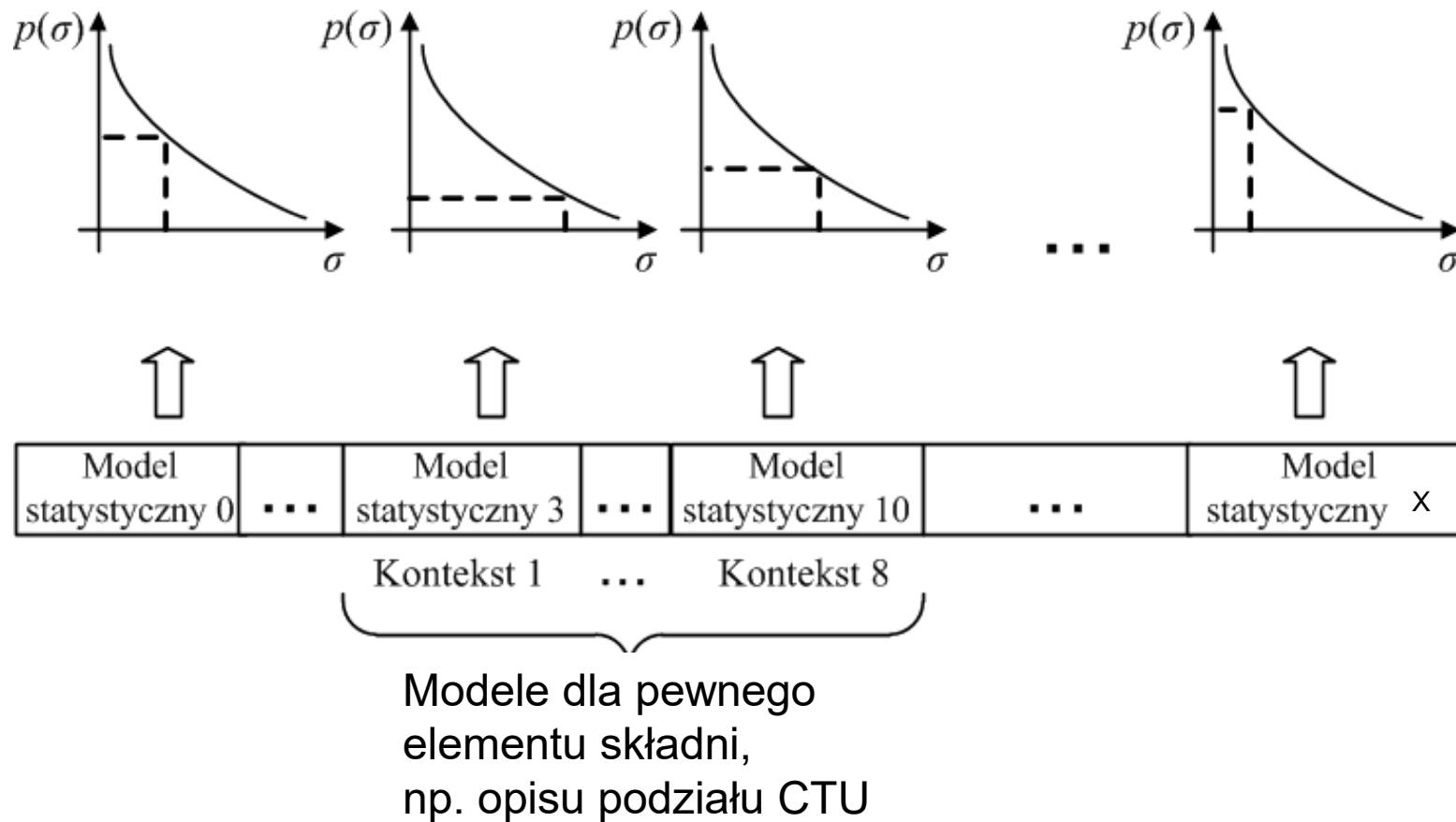


symbols in CABAC:

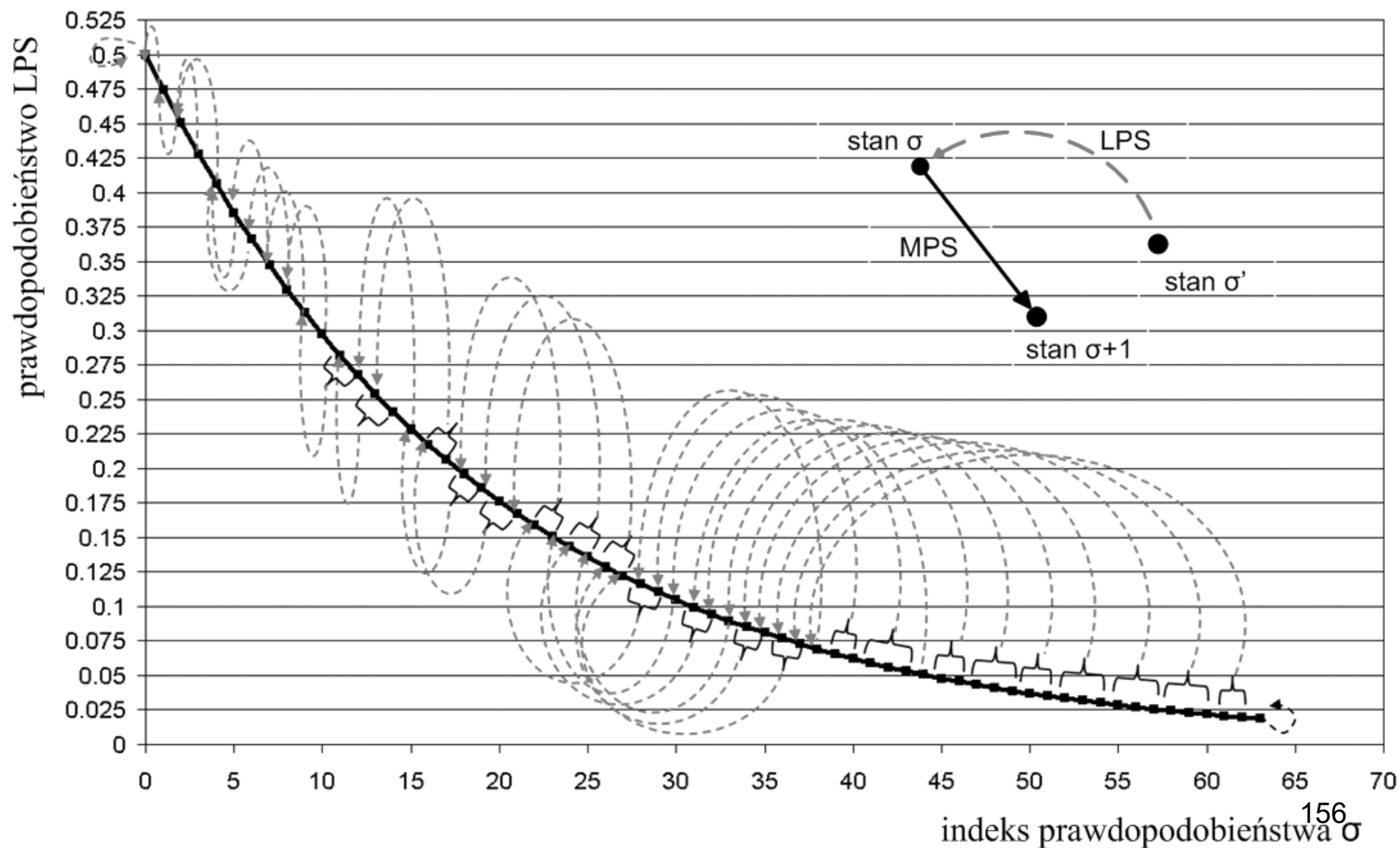
LPS – least probable symbol

MPS – most probable symbol

Modele prawdopodobieństwa



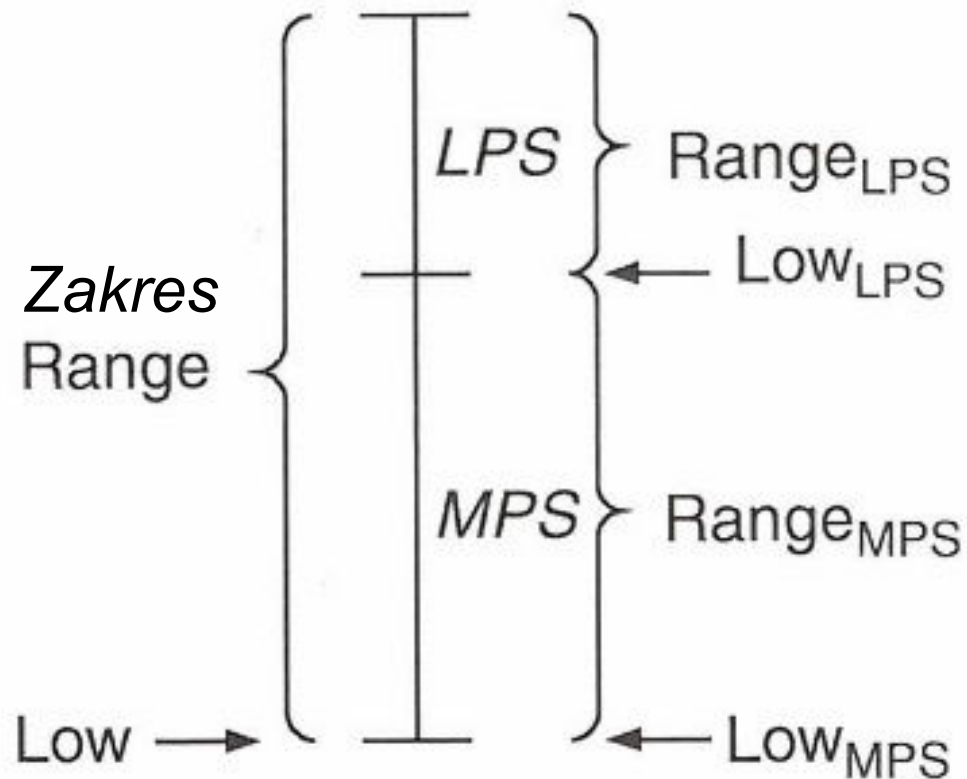
Automat wyznaczania prawdopodobieństwa dla LPS – osobny dla każdego symbolu i kontekstu



Koder binarny arytmetyczny

- Przeskalowywanie zakresu, gdy staje się mniejszy od $\frac{1}{4}$ maksymalnego zakresu
- Bity pojawiające się na skutek przeskalowania zakresu stanowią kod.
- Wartości prawdopodobieństw są skwantowane

Koder binarny arytmetyczny



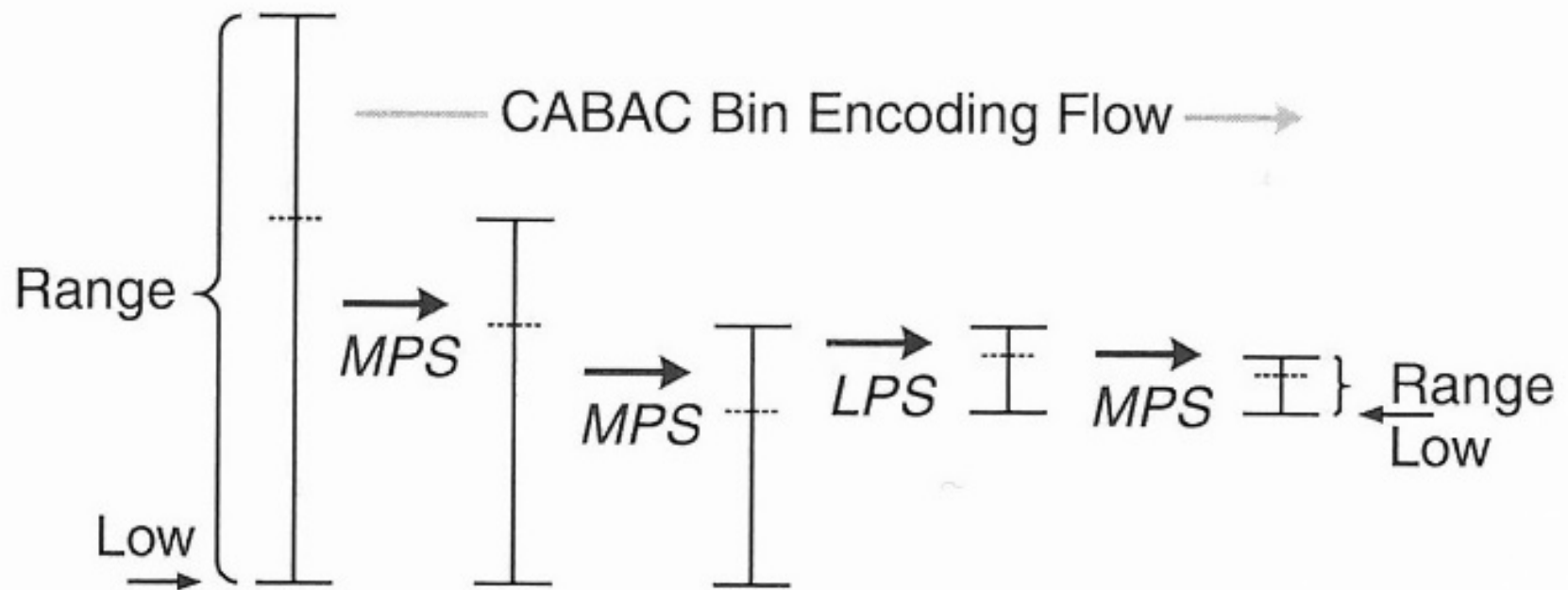
LPS – Least Probable Symbol

– *Najmniej prawdopodobny symbol*

MPS – Most Probable Symbol

– *Najbardziej prawdopodobny symbol*

Koder binarny arytmetyczny



8.13. Sterowanie pracą kodera

Tryb stałej jakości $\Rightarrow$ zmienna prędkość bitowa

Minimalizacja całkowitej objętości pliku

Prędkość bitowa nie jest krytyczna

Przykład: nagrywanie płyt DVD

B-ISDN z UPC (*Usage Parameter Control*)

Tryb stałej prędkości bitowej $\Rightarrow$ zmienna jakość

Krytyczna jest dostępna przepustowość kanału, którą trzeba wykorzystywać

Ograniczenie: dopuszczalne opóźnienie lub dopuszczalna wielkość bufora

Przykład: telewizja cyfrowa

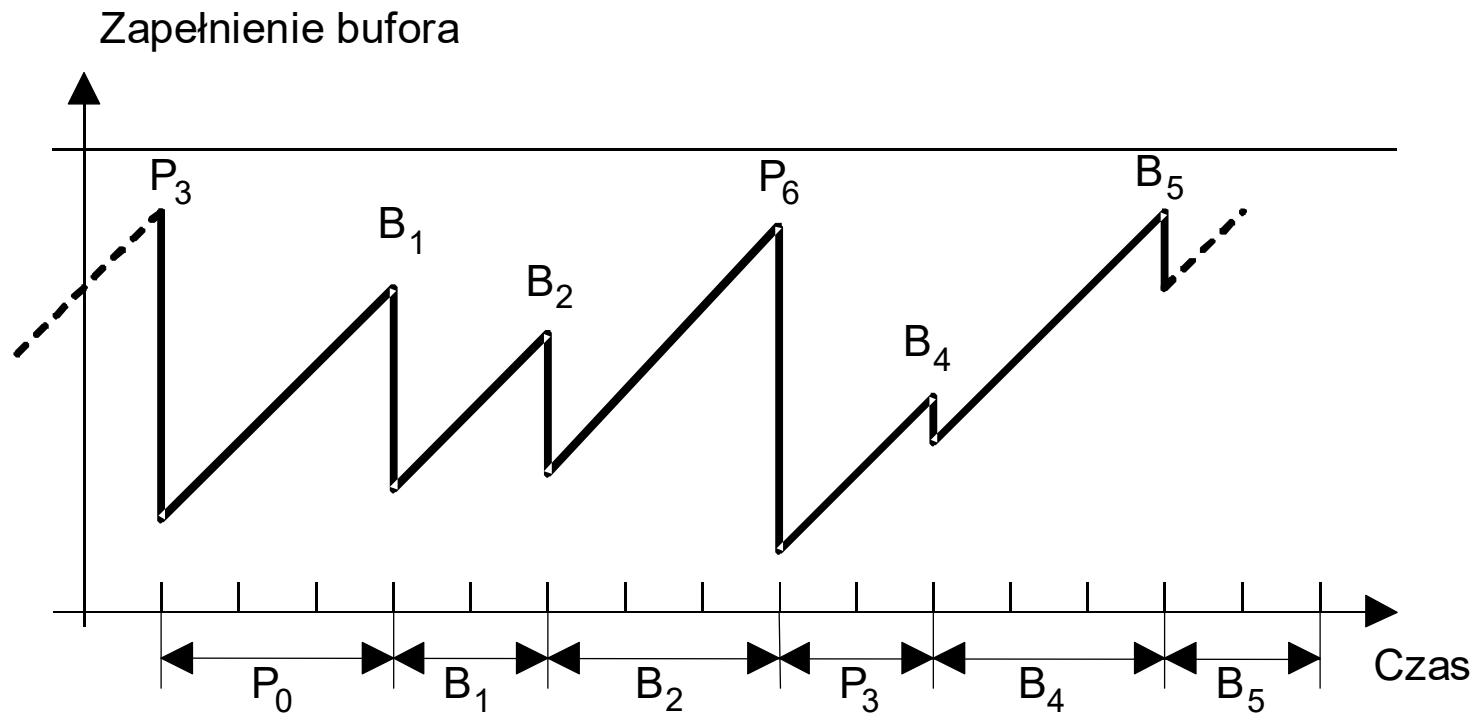
Sterowanie pracą kodera hybrydowego (c.d.)

Inne sformułowanie :

- **Tryb stałej prędkości transmisji**
CBR (*constant bitrate*)
- jakość zmienna .
- **Tryb zmiennej prędkości transmisji**
VBR (*variable bitrate*)
- jakość stała lub prawie stała.

Buforowanie

Wielkość bufora jest określana w nagłówku sekwencji



VBV – video buffering verifier

Mechanizmy regulacji

- Dobór wartości współczynnika Q skali kwantyzatorów
- Pomijanie obrazów

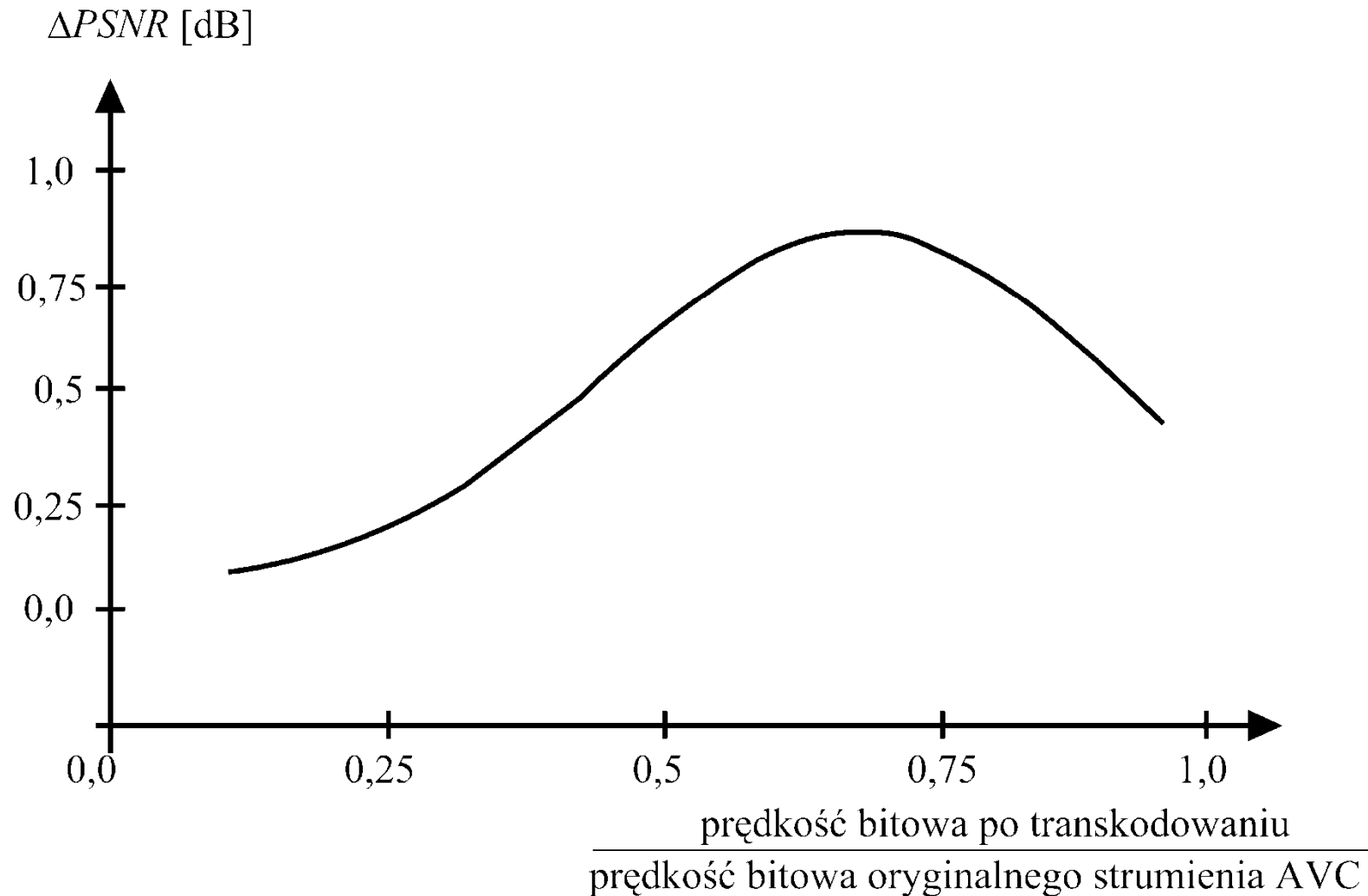
Sterowanie globalne i lokalne

- Sterowanie globalne
- Sterowanie lokalne

Hipotetyczny dekodery modelowy ***Hypothetical Reference Decoder (HDR)***

Wielokrotne kodowanie tego samego materiału

- Utrata jakości



8.14. Przesyłanie strumieni wizyjnych

- **Wymiana plików: Kontenery:**
np. **AVI (Audio Video Interleave)**.
 - Paczki danych (*chunks*) ze fragmentami strumieni wizji i fonii.
 - Strumienie uzupełnione dodatkowymi znacznikami i danymi porządkowymi (np. o ilości danych).
 - Format kontenera niezależny od technik kompresji wizji i fonii.
 - Mechanizm synchronizacji wizji i fonii.
- **Strumieniowanie**

Strumieniowanie wizji

- **Strumień transportowy MPEG-1:**
 - Podział strumieni wizji i fonii na krótkie pakiety o stałej długości (184 bajty + 4 bajty nagłówek).
 - Mechanizmy synchronizacji i korekcji błędów transmisji.
- **Dynamiczne adaptacyjne strumieniowanie za pomocą HTTP** (*Hypertext Transfer Protocol*)
Dynamic Adaptive Streaming over HTTP (DASH):
 - Strumieniowanie adaptacyjne

8.15. Efektywność technik MPEG-2, AVC, HEVC, VVC

Kompresja wizji

Wymagana szybkość (prędkość) bitowa.

Założenia dla dalszych rozważań:

1. **Omawiane są wartości orientacyjne szybkości transmisji dla dobrej (telewizyjnej) jakości obrazu** dekodowanego w odbiorniku.
2. Wartości podano dla **wymagających treści** (często szybki ruch kamery i/lub obiektów). Często mniejsze szybkości bitowe są używane (nawet stanowiące czasem tylko 15% podanych wartości dla prostych treści, np. dla programów publicystycznych).
3. Założono **zastosowanie dobrych koderów** w pełni wykorzystujących możliwości danej normy.

Efektywność kompresji

Generacja systemu tv	Format wizji (lumi)	Częstotliwość obrazów [Hz]	Szybkość bitowa [Mb/s]			
			MPEG-2	AVC	HEVC	VVC
SDTV	720 × 576	25	4	2	1	0,6
HDTV	1920 × 1080	25	16	8	4	2
UHDTV	3840 × 2160	50	--	32	16	8

Dla 8-bitowych (UHDTV: 10-bitowych) próbek lumi i chromy, dla próbkowania chromy 4:2:0.

Uwaga: 1. **Wartości orientacyjne** dla **dobrej (telewizyjnej) jakości obrazu** dekodowanego w odbiorniku. Wartości podano dla **wymagających treści** (często szybki ruch kamery i/lub obiektów). Często mniejsze szybkości bitowe są używane (nawet stanowiące czasem tylko 15% podanych wartości dla prostych treści , np. dla programów publicystycznych).

2. Założono **zastosowanie dobrych koderów** w pełni wykorzystujących możliwości danej normy.

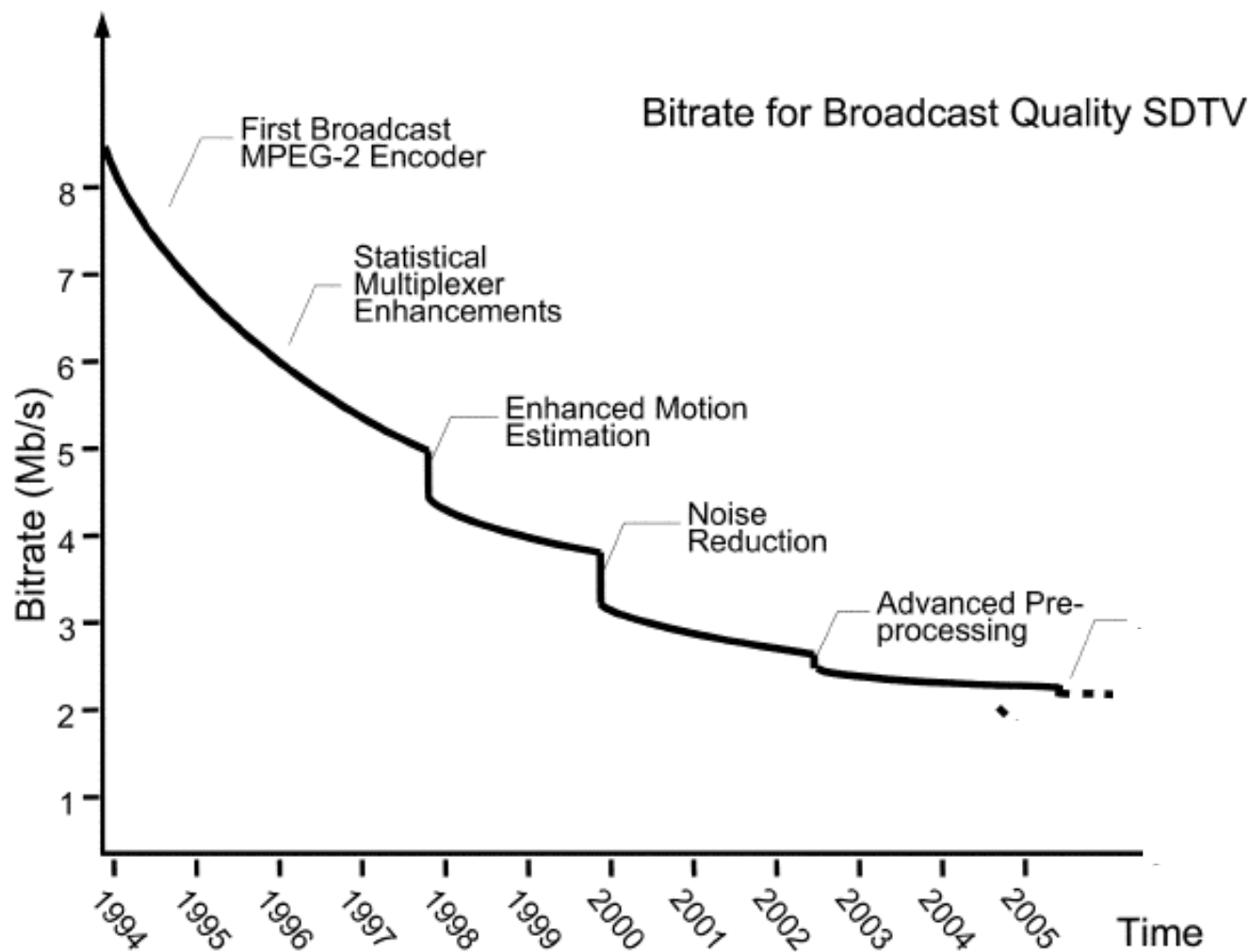
Efektywność kompresji

Efektywność kompresji dla charakterystycznych zastosowań w telewizji
(f_o oznacza częstotliwość obrazu)

Nazwa normy	Zastosowanie	Orientacyjna prędkość bitowa [Mb/s]	Przybliżona średnia liczba bitów przypadająca na punkt obrazu kolorowego
MPEG-2	SDTV ($720 \times 576, f_o = 25 \text{ Hz}$)	4	0,40
AVC	HDTV ($1920 \times 1080, f_o = 25 \text{ Hz}$)	8	0,15
HEVC	UHDTV ($3840 \times 2160, f_o = 50 \text{ Hz}$)	16	0,04
VVC		8	0,02

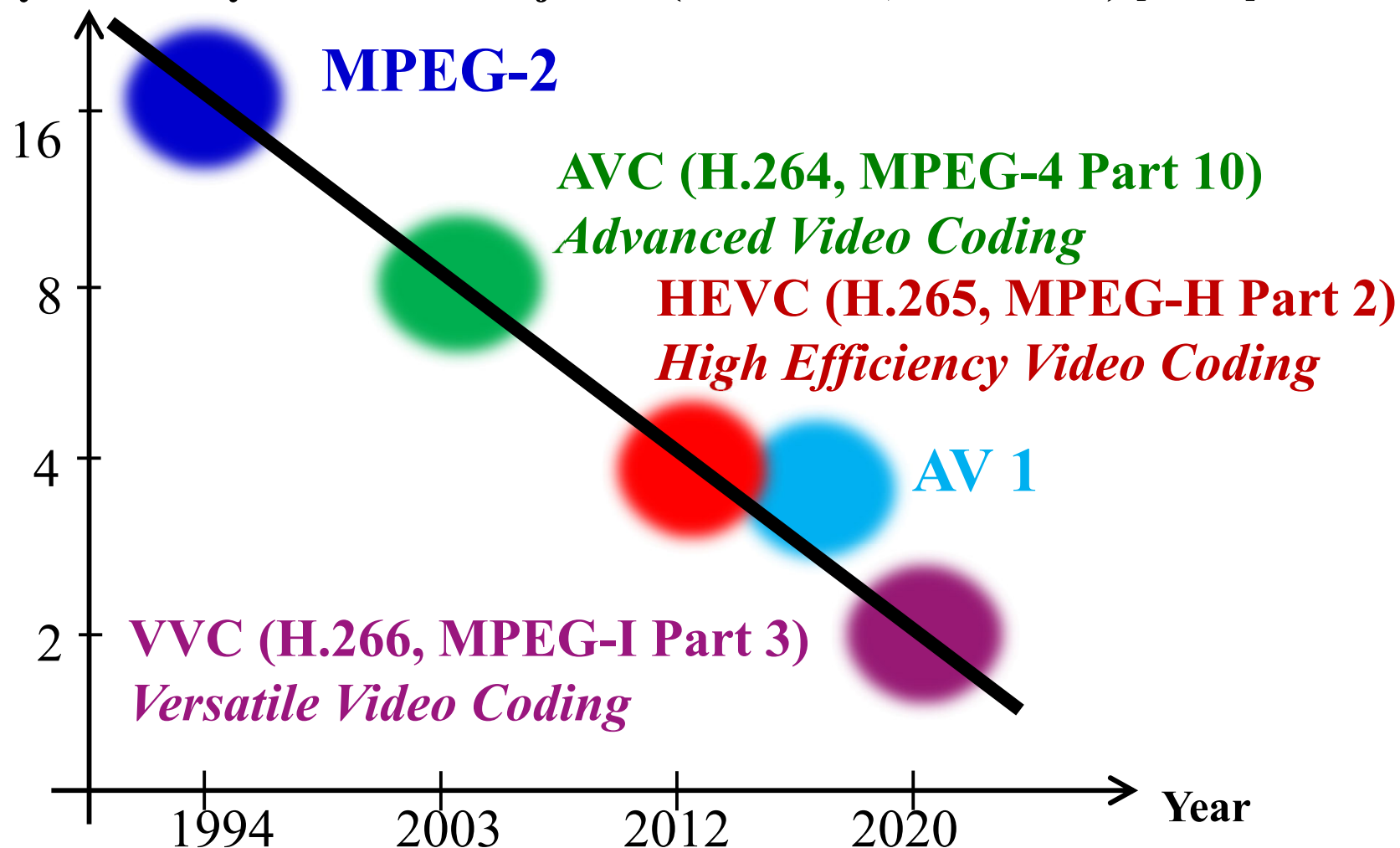
Uwaga: W praktyce stosowane są często (znacznie) mniejsze prędkości bitowe. Założono **zastosowanie dobrych koderów** w pełni wykorzystujących możliwości danej normy.

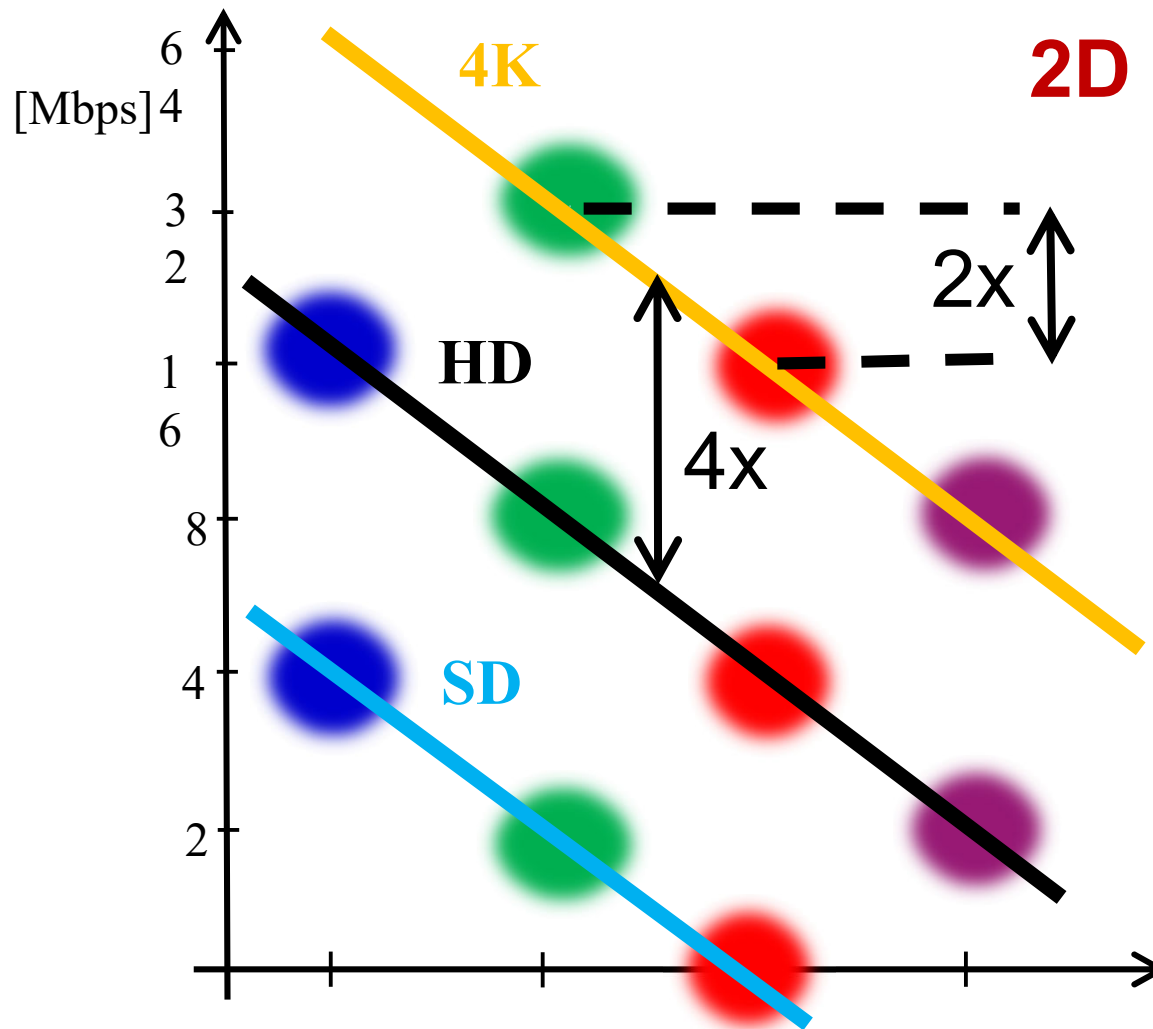
Przykład historyczny: SDTV o jakości rozsiewczej – kodowanie MPEG-2



Kompresja wizji

Przybliżona szybkość transmisji **HD** (1920×1080, 25 frame/s) [Mb/s]





2D video coding

MPEG-2

(H.262)

AVC (H.264,

MPEG-4 Part 10)

Advanced Video Coding

HEVC

(H.265, MPEG-H Part

2)

High Efficiency Video Coding

VVC

Versatile Video

Coding

(H.266,

MPEG-I part 3)

Szybkość bitowa

$$B \approx A \cdot V \quad [\text{Mbps}]$$

gdzie A jest współczynnikiem zależnym od techniki kompresji,

$A = 4$ dla MPEG-2,

$A = 2$ dla AVC,

$A = 1$ dla HEVC,

$A = 0.5$ dla VVC

a V jest współczynnikiem zależnym od formatu obrazu,

$V=1$ dla SDTV (720×576, 25i),

$V=4$ dla HDTV (1920×1080, 25i),

$V=16$ dla UHD TV (3840×2160, 50p),

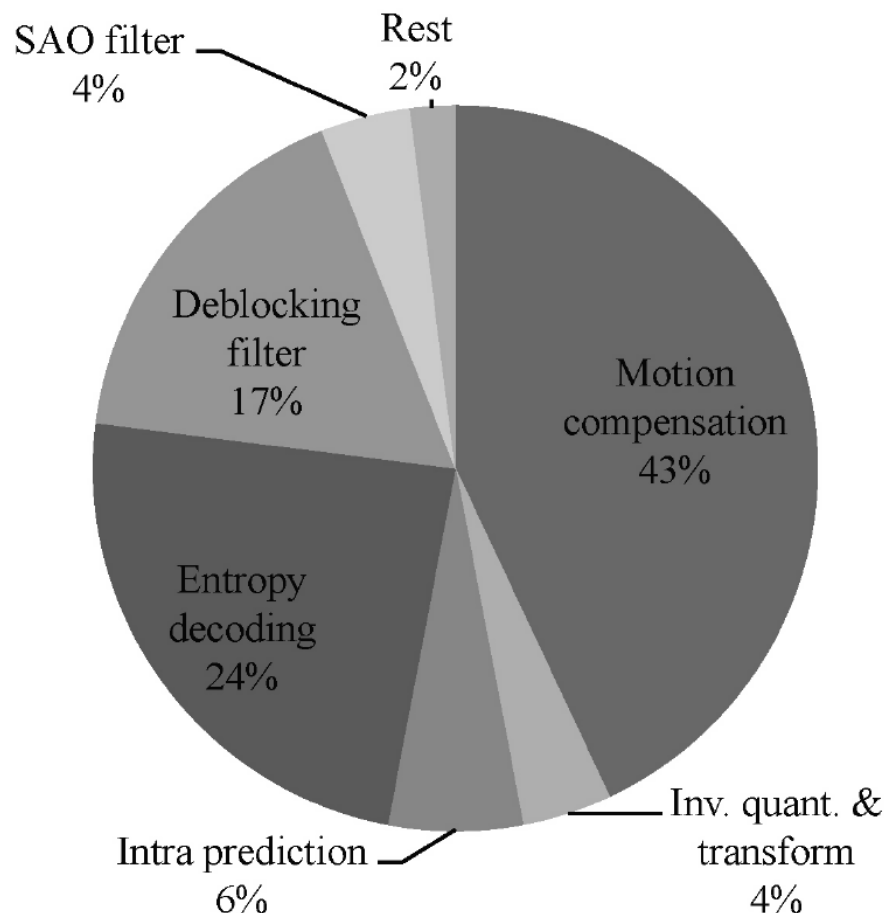
Liczba wariantów kodowania bloku 16 x 16 próbek obrazu

- **MPEG-2** MP@HL 4 warianty
- **AVC** HP@L4 $\sim 10^6$ wariantów
- **HEVC** MP@L4 $\sim 10^9$ wariantów

Dekoder - złożoność

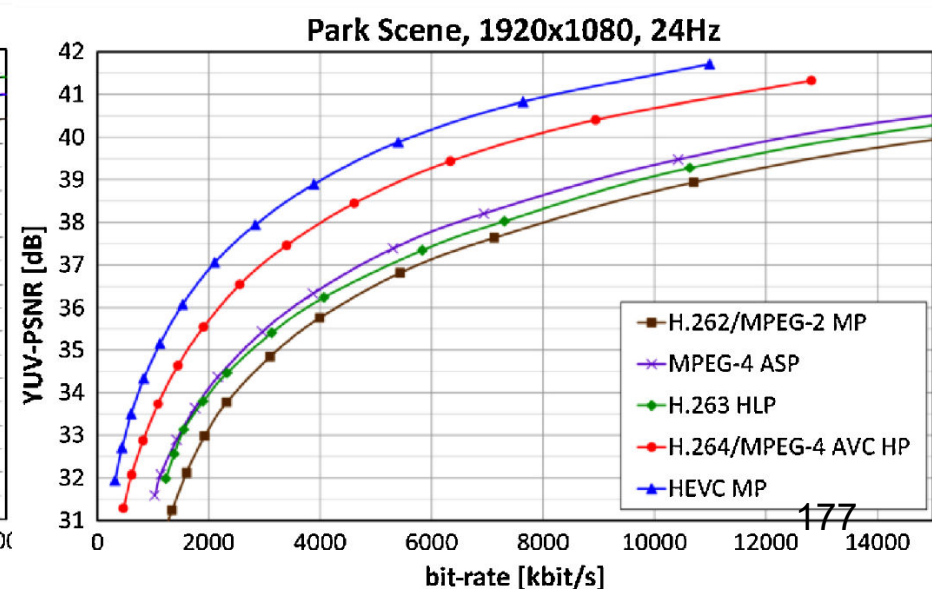
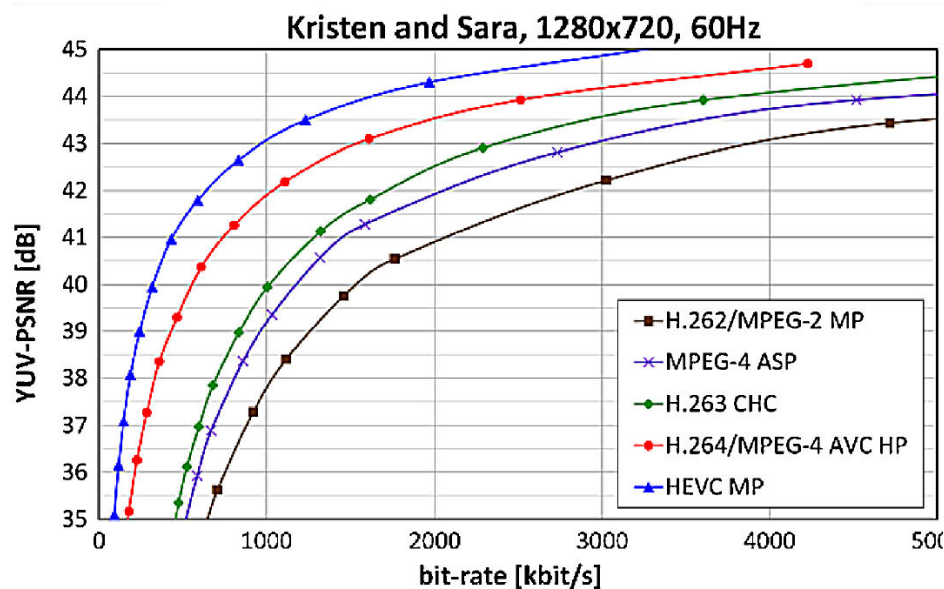
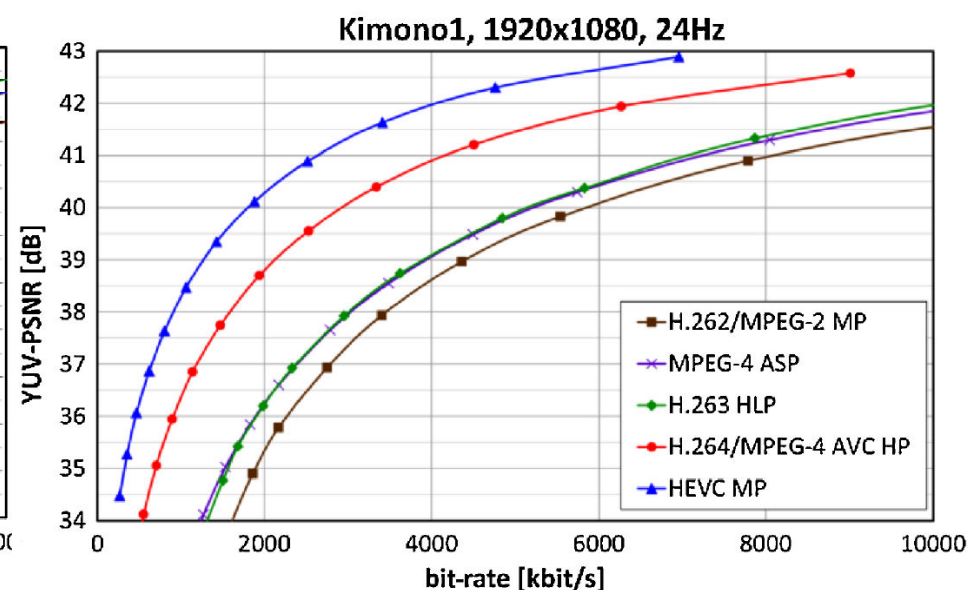
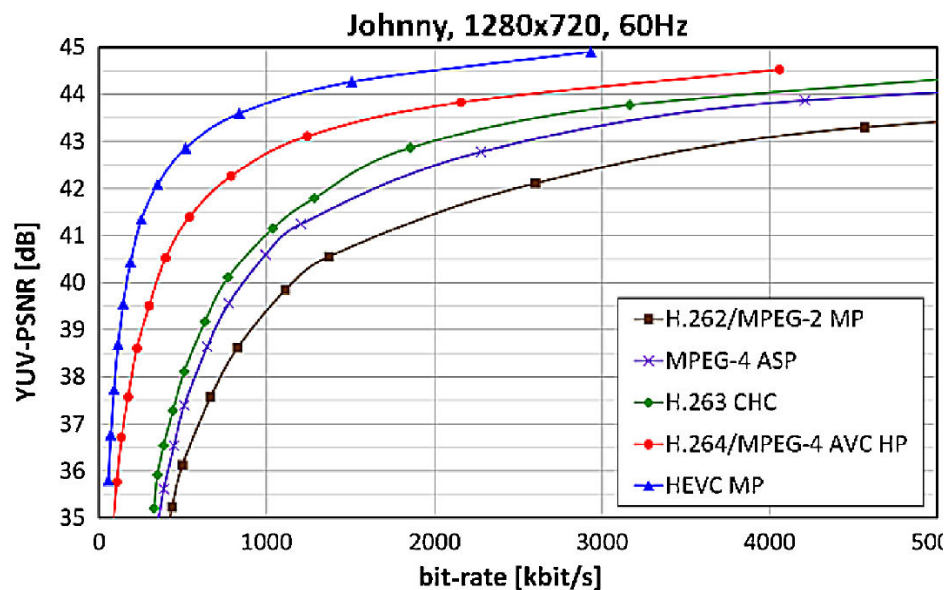
- Zoptymalizowany dekodery HEVC ma podobne wymagania obliczeniowe / sprzętowe do zoptymalizowanego dekodera AVC !!

Random Access (ARM)



Porównanie jakości metodą obiektywną

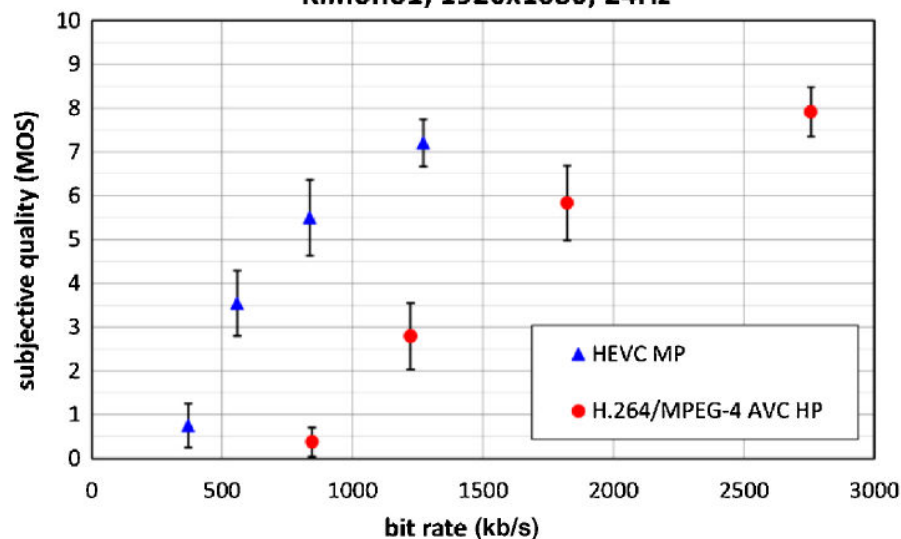
Objective performance comparison



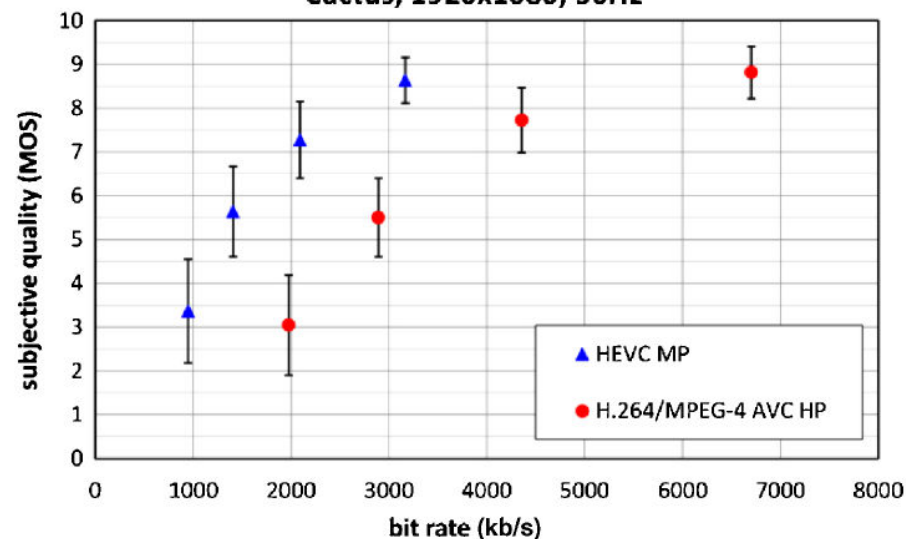
Porównanie jakości metodą subiektywną

Subjective performance comparison

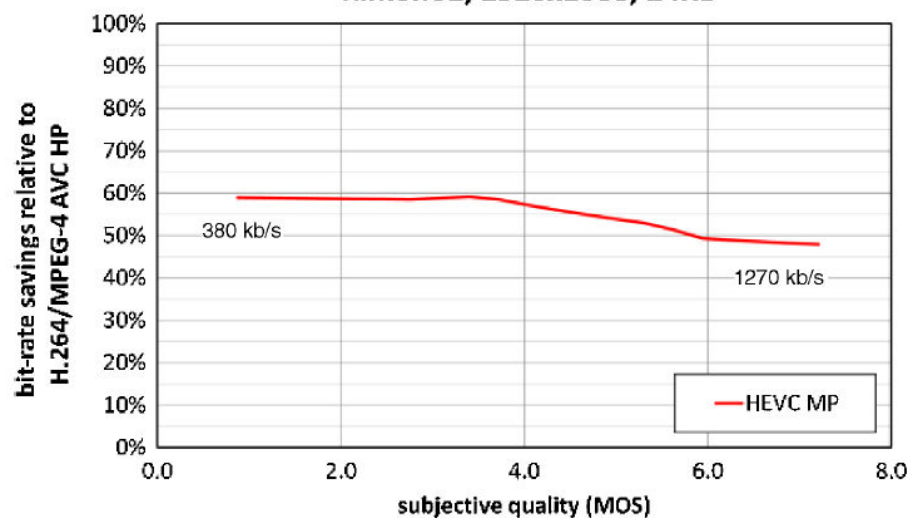
Kimono1, 1920x1080, 24Hz



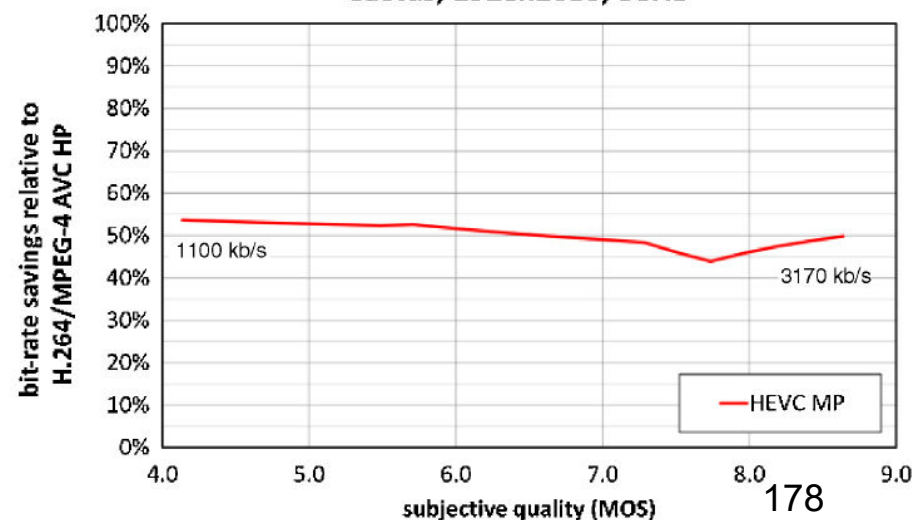
Cactus, 1920x1080, 50Hz



Kimono1, 1920x1080, 24Hz



Cactus, 1920x1080, 50Hz



HEVC w dozorze wizyjnym

a



b



c



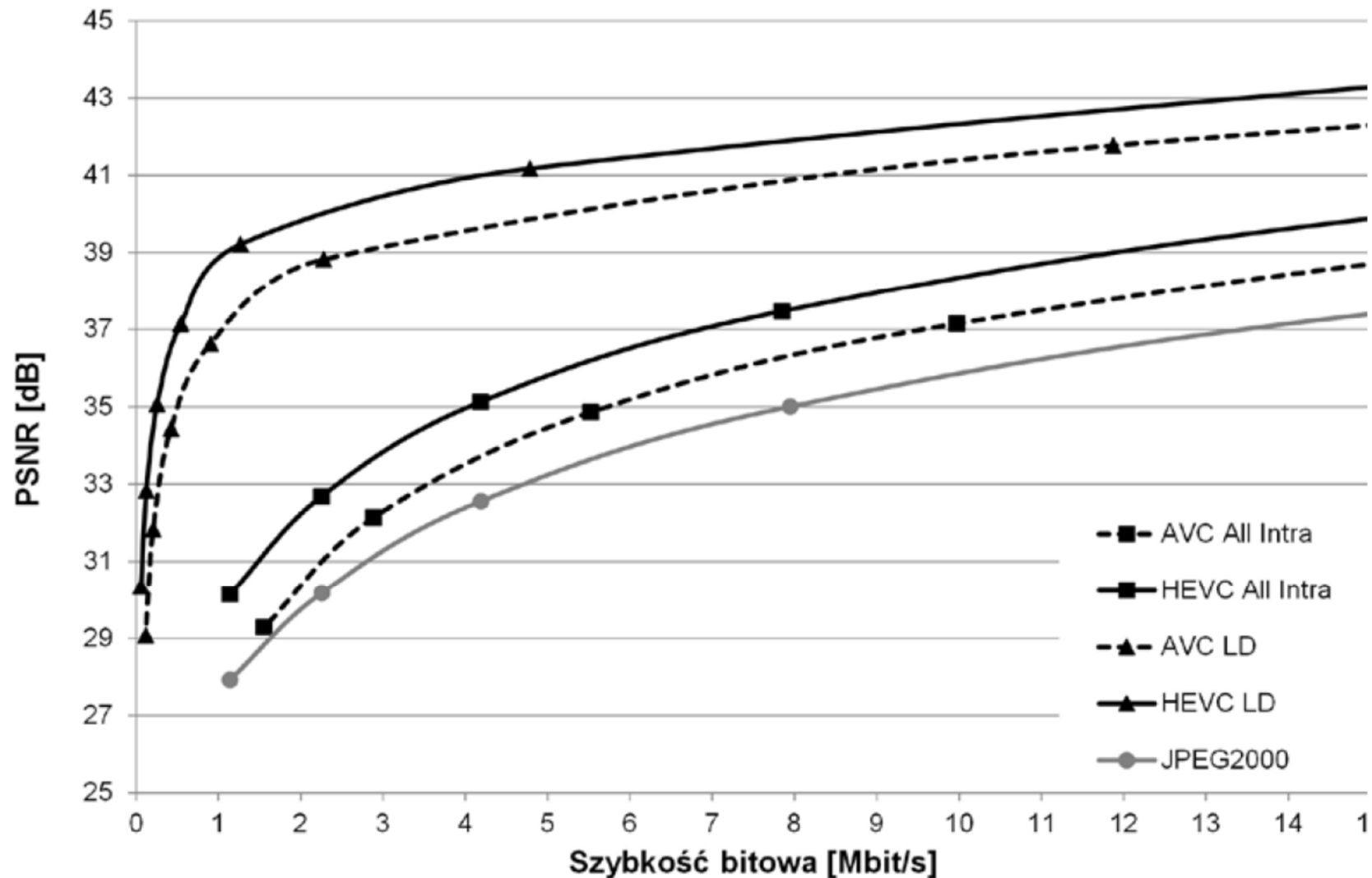
d



e

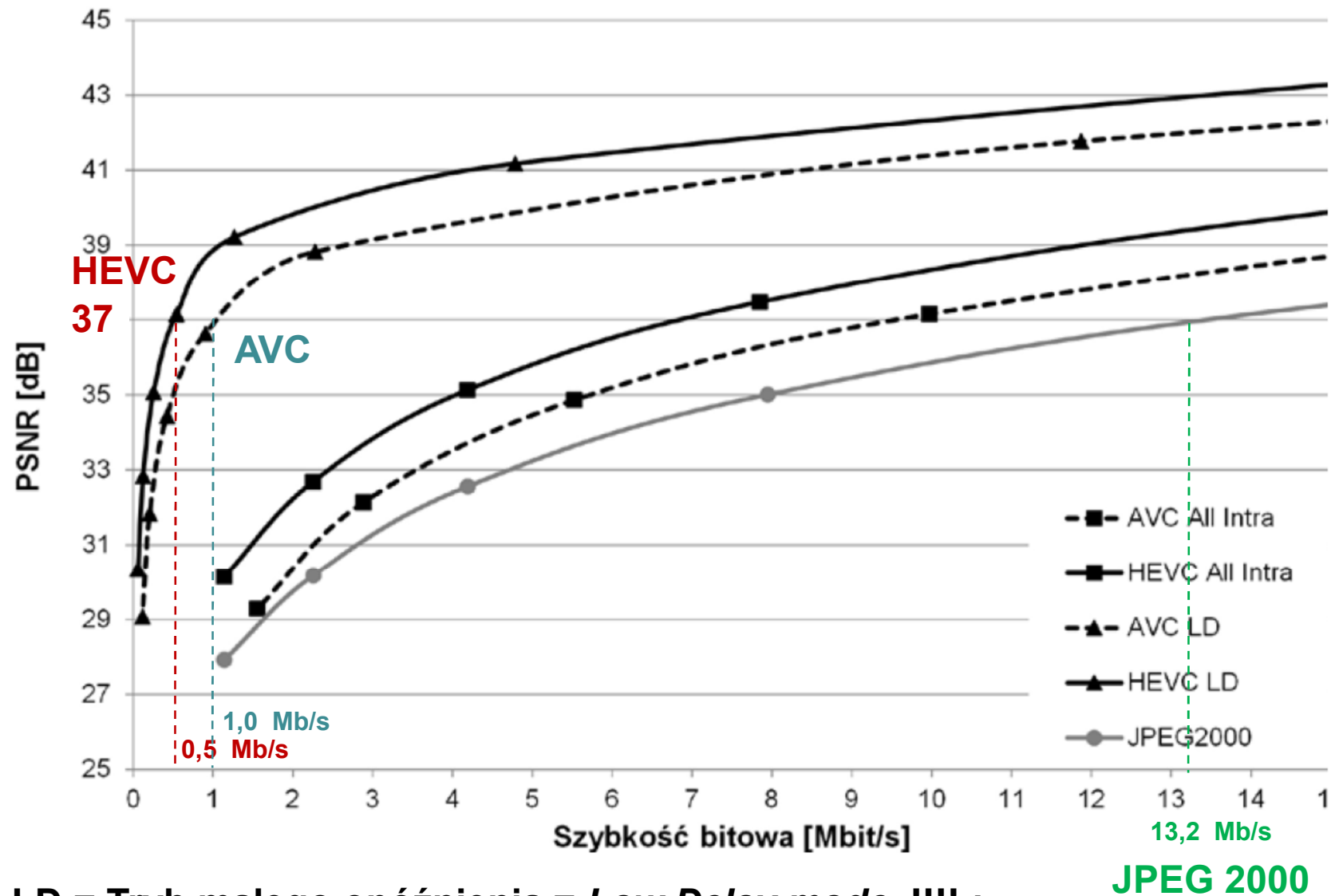


HD (1920 x 1080), 25 fps, w. kolejnoliniowe



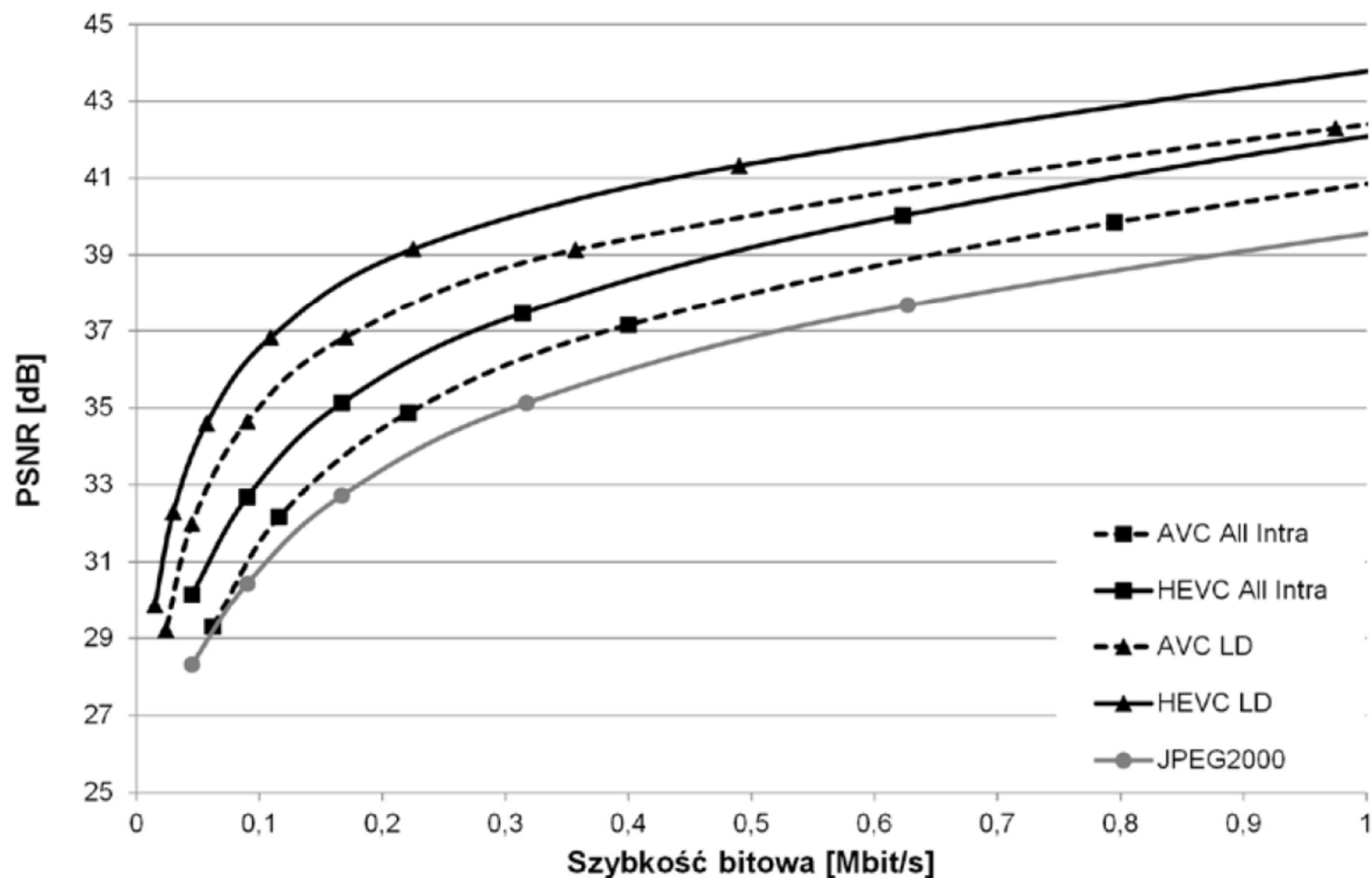
LD = Tryb małego opóźnienia = *Low Delay mode*:
No bidirectional prediction, Intra period = 2 seconds.

HD (1920 x 1080), 25 fps, w. kolejnoliniowe



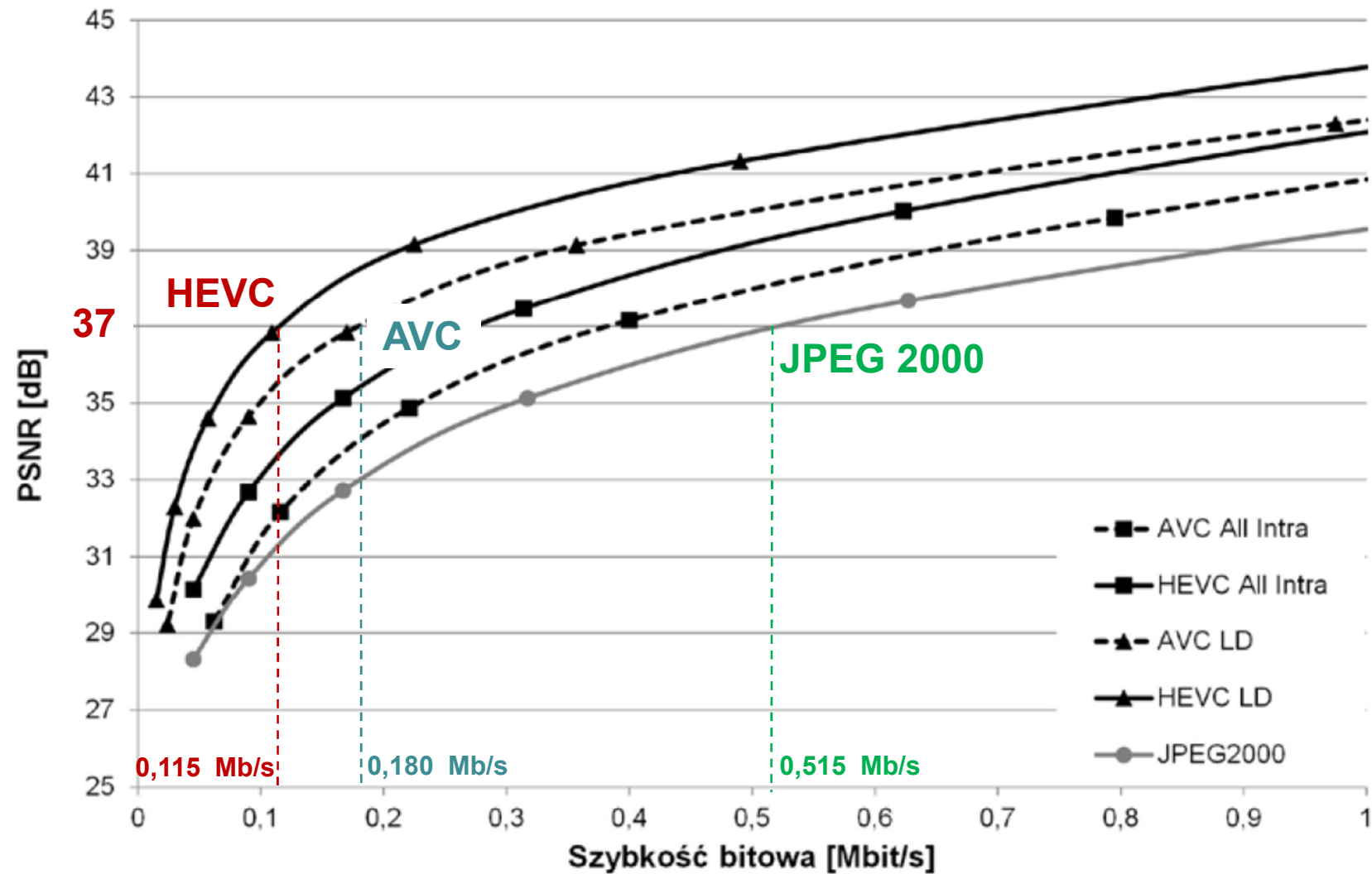
LD = Tryb małego opóźnienia = *Low Delay mode* !!!! :
No bidirectional prediction, Intra period = 2 seconds.

HD (1920 x 1080), 1 fps, w. kolejnoliniowe



LD = Tryb małego opóźnienia = *Low Delay mode*:
No bidirectional prediction, Intra period = 2 seconds.

HD (1920 x 1080), 1 fps, w. kolejnoliniowe



LD = Tryb małego opóźnienia = *Low Delay mode*:
No bidirectional prediction, Intra period = 2 seconds.

Efektywność technik
MPEG-2, AVC, HEVC, VVC
Ciąg dalszy

HEVC 1,3 Mb/s



AVC (MPEG-4 cz. 10, H.264) 1,3 Mb/s



AVC (MPEG-4 cz. 10, H.264) 1,3 Mb/s



MPEG-2 1,3 Mb/s



MPEG-2 1,3 Mb/s



MPEG-2 5,2 Mb/s



MPEG-2 5,2 Mb/s



HEVC 1,3 Mb/s



HEVC: sekwencja „Bus”, ~1,0 Mb/s



AVC: sekwencja „Bus”, ~2,0 Mb/s



MPEG-2: sekwencja „Bus”, ~4,0 Mb/s



MPEG-2: sekwencja „Bus”, ~1,0 Mb/s



MPEG-2: sekwencja „Bus”, ~1,0 Mb/s



AVC: sekwencja „Bus”, ~1,0 Mb/s



AVC: sekwencja „Bus”, ~1,0 Mb/s



HEVC: sekwencja „Bus”, ~1,0 Mb/s



HEVC: sekwencja „Cheer” SD, ~1,5 Mb/s



AVC: sekwencja „Cheer”, ~3,0 Mb/s



MPEG-2: sekwencja „Cheer”, ~6,0 Mb/s



AVC: sekwencja „Cheer”, ~1,5 Mb/s



AVC: sekwencja „Cheer”, ~1,5 Mb/s



MPEG-2: sekwencja „Cheer”, ~1,5 Mb/s



MPEG-2: sekwencja „Cheer”, ~1,5 Mb/s

