

Multimedia

MARIUSZ SZWOCH

KATEDRA INTELIGENTNYCH SYSTEMÓW INTERAKTYWNYCH WETI, PG

1. Multimedia

"Integracja wielu różnorodnych **mediów** (telewizji, techniki audio i wideo, informatyki, teletransmisji) na pewnej wspólnej bazie, która może stanowić **komputer**, specjalny odtwarzacz płyt kompaktowych lub przystawka do odbiornika telewizyjnego; także nazwa systemu umożliwiającego swobodną, **interaktywną** wymianę informacji w postaci tekstu, grafiki, obrazu (nieruchomego i ruchomego), dźwięku (mowy, muzyki) itp. między różnymi jego elementami. [...]"

Nowa encyklopedia powszechna PWN, Warszawa 1997

Multimedia - definicje (2)

- **Multimedia** oznaczają komunikację wykorzystującą **wiele form przekazu**, postrzeganą przez odbiorcę **wieloma kanałami percepcji**
- **Interaktywne multimedia** pozwalają użytkownikowi na kontrolę treści i formy docierających do niego informacji
- **Hipermedia** posiadają strukturę połączonych elementów (hiperłącza), po których użytkownik może się swobodnie poruszać

Media

- Składniki multimedialnych:

- tekst
- obraz
 - statyczny (grafika/zdjęcia)
 - ruchomy (film/animacja)
- dźwięk
 - muzyka
 - mowa/śpiew



Historia multimedialnych

- Multimedialne pakiety pomocy naukowych
- Wczesne komputery multimedialne (audio-video computer)
- Komputerowe systemy multimedialne 1986 r.
 - kolorowy monitor
 - karta graficzna obsługująca kolor
 - karta dźwiękowa
 - głośniki stereo lub wyjście audio
 - stacja CD-ROM z możliwością odtwarzania:
płyt audio, płyt z danymi, płyt mieszanych (*mixed-mode*).
- Współczesne komputery multimedialne
 - pozyskiwanie, przetwarzanie i odtwarzanie danych graficznych, dźwiękowych, animacji i wideo

Multimedia PC Standards

- Grupa *Multimedia PC Marketing Council* (Microsoft, AT&T, NEC, CreativeLabs, Fujitsu, Philips, ..)
- MPC I (1991):
 - 386 SX (16MHz), 4MB RAM, 30 MB HDD, audio (8bit+MIDI), VGA (640x480x8), CD-ROM (150 kB/s, 600 ms)
- MPC II (1994):
 - 486 SX (25MHz), 8MB RAM, 160 MB HDD, audio (16bit+MIDI), VGA (640x480x16), CD-ROM (300 kB/s, 400ms)
- MPC III:
 - Pentium (75MHz), 16MB RAM, 500 MB HDD, audio (16bit+MIDI), SVGA+MPEG1 (800x600x8), CD-ROM (600 kB/s, 280 ms)
- MPC 2006:
 - Pentium IV (3GHz), 512MB RAM, 120 GB HDD, audio (16bit+MIDI+5.1+surround), XSVGA (32 MB, 1280x1024x24), DVD±RW (5 MB/s, 150 ms)

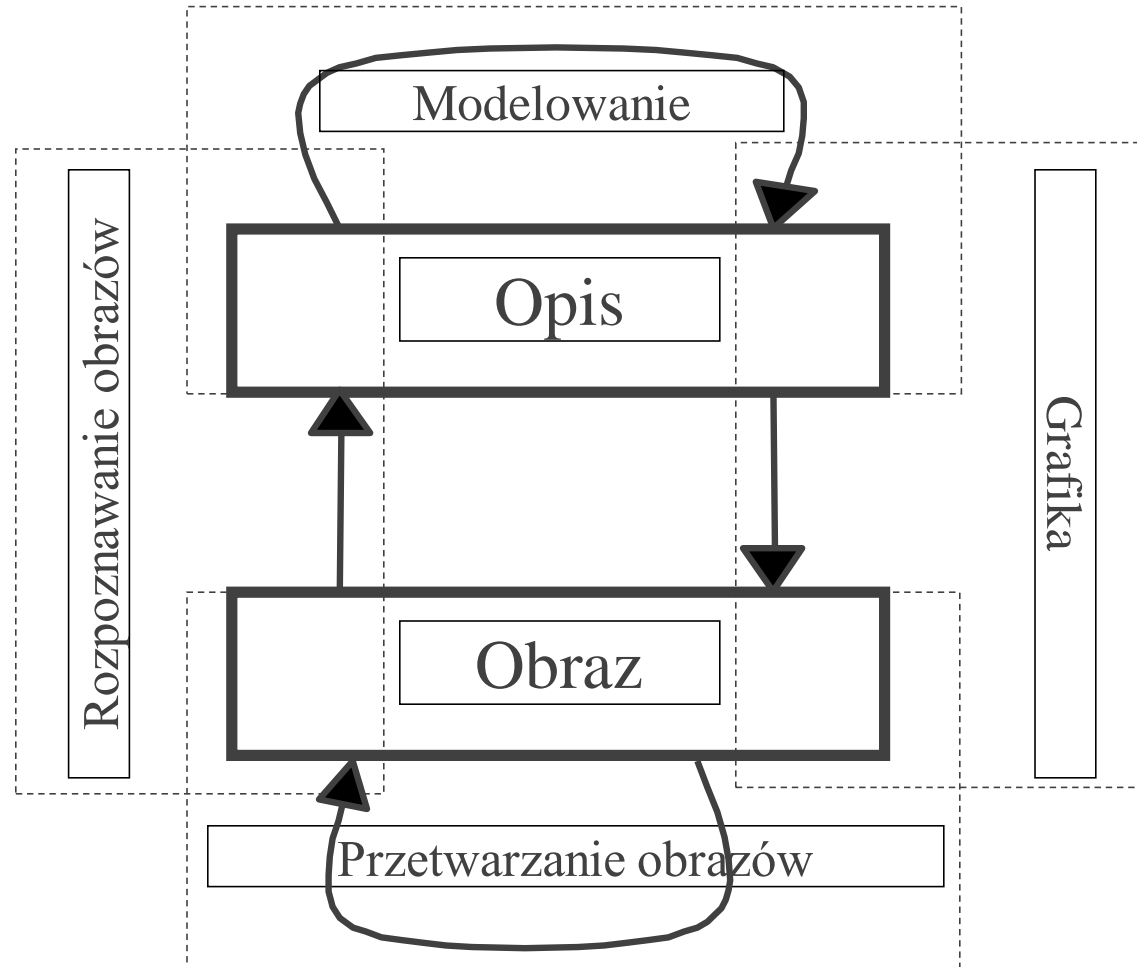
Przykładowe zastosowania multimediiów

- Rozrywka (gry, karaoke)
- Symulacje komputerowe
- Edukacja (również nauczanie na odległość)
- Hipermedia (aplikacje Internetowe)
- Prezentacje multimedialne
- Komunikacja (telekonferencje, poczta multimedialna)
- Hobby (edycja dźwięku, wideo, fotografii)
- Pomoc osobom niepełnosprawnym
- Systemy mobilne

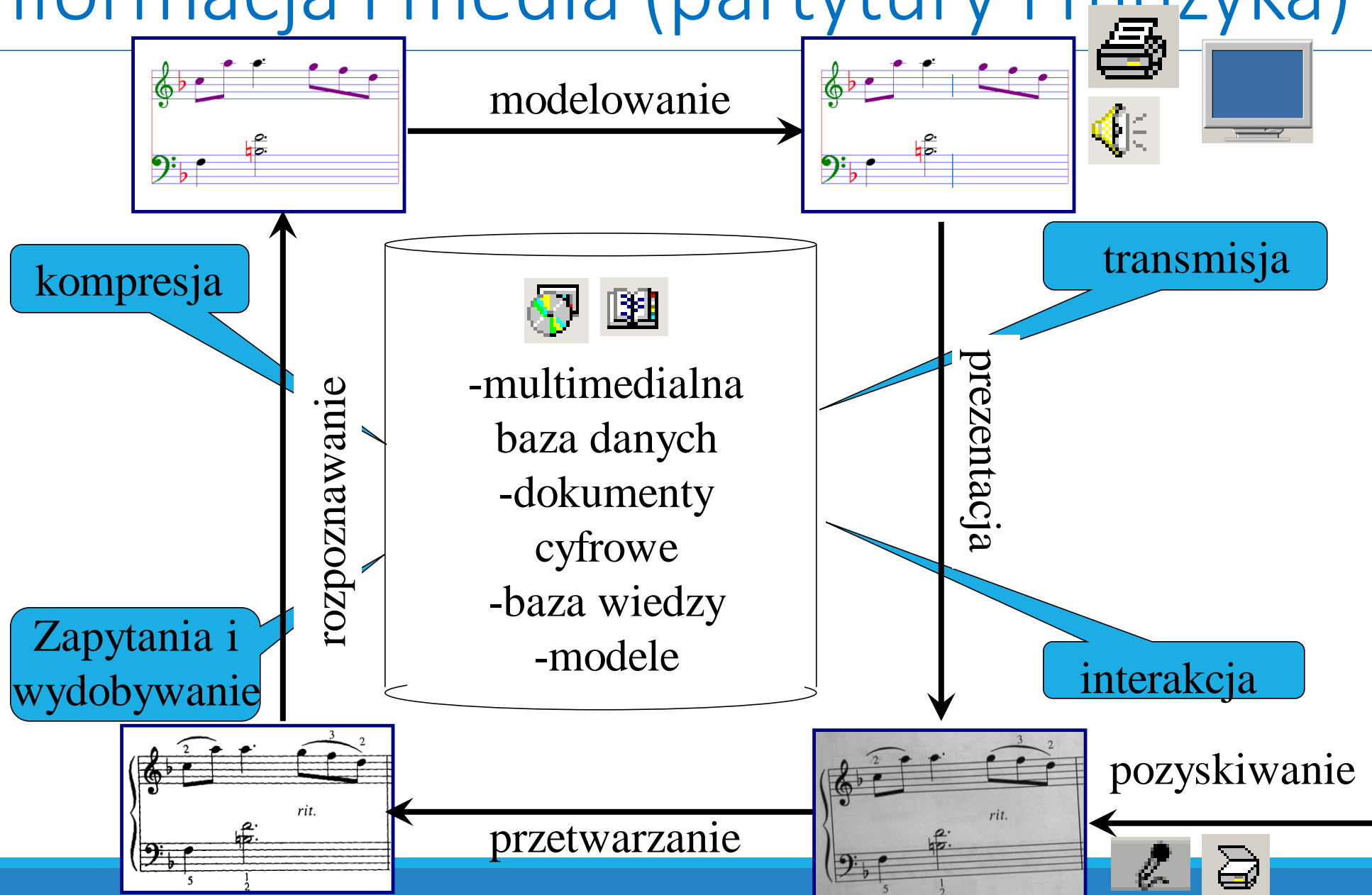
Kierunki rozwoju systemów multimedialnych

- Systemy teletransmisji (*Internet, Video-On-Demand, TV & radio broadcasting*)
- Wirtualna, rozszerzona i mieszana rzeczywistość
- Inteligentne domy i domowe centra rozrywki
- Nowe rodzaje interfejsów
 - rozpoznawanie gestów i ruchu ciała NUI (ang. *natural user interface*)
 - rozpoznawanie i synteza mowy VUI (ang. *voice UI*)
 - rozpoznawanie pisma odręcznego
 - rozpoznawanie ruchów gałki ocznej

Informacja i dane multimedialne (obraz)



Informacja i media (partytury i muzyka)

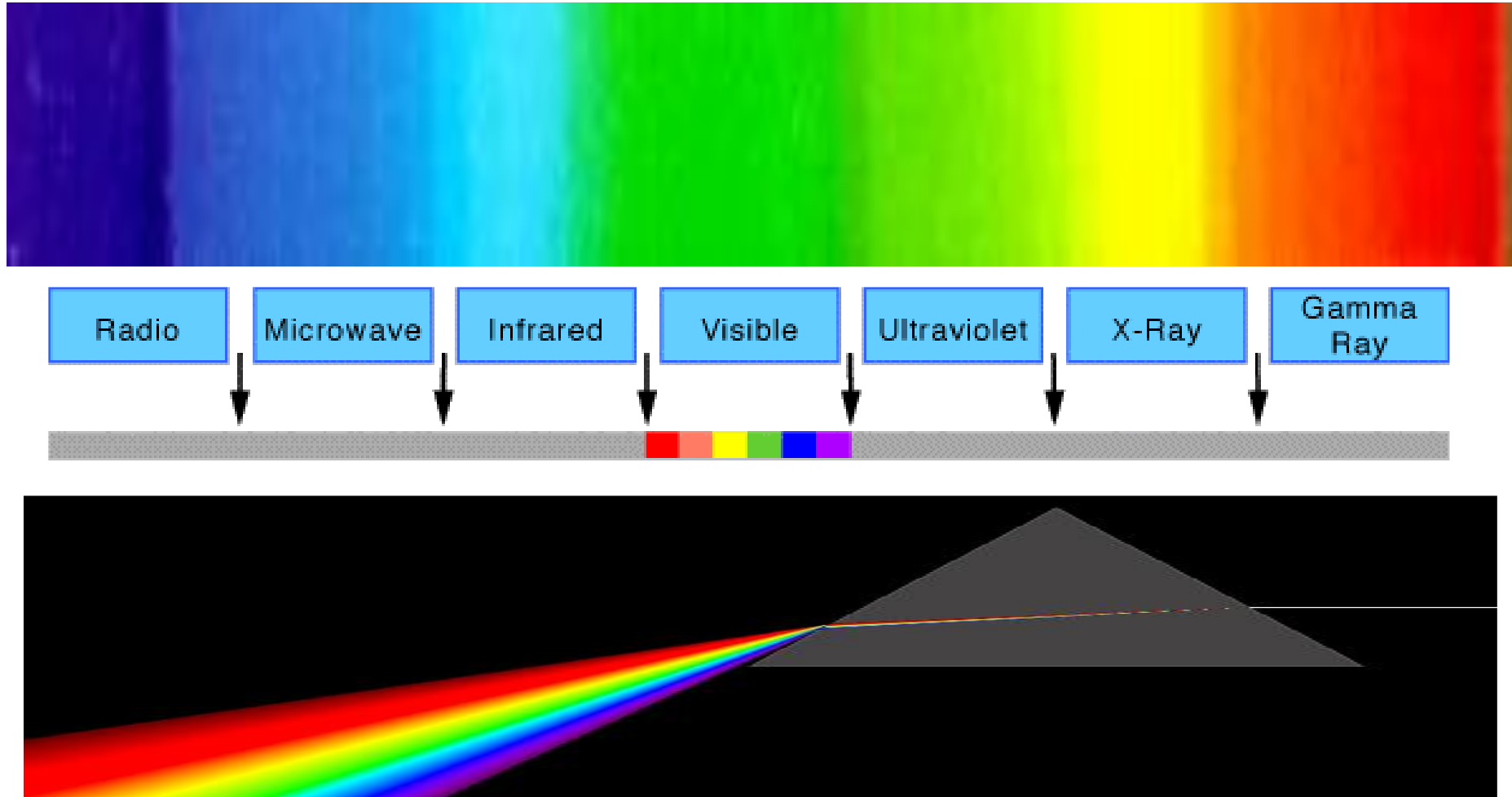


1.1. Nośniki informacji

- **Nośnik informacji** - medium umożliwiające dostarczanie człowiekowi informacji o świecie zewnętrznym
- Rodzaje nośników informacji (zmysły i odczucia)
 - **światło widzialne** (wzrok)
 - **fala akustyczna - dźwięk** (słuch)
 - związki chemiczne (zapach, smak)
 - inne:
 - oddziaływanie fizyczne (dotyk, równowaga)
 - ładunek i prąd elektryczny (ból)
 - promieniowanie podczerwone (ciepło)

1.1.1. Światło

- *Światło* - część widzialna promieniowania elektromagnetycznego.



Barwa

- **Barwa** - wrażenie psychofizyczne odczuwane za pośrednictwem zmysłu wzroku pod wpływem światła o określonym składzie widmowym.

Encyklopedia Powszechna PWN

- **Barwa prosta** (widmowa, spektralna, monochromatyczna) – barwa wywołana promieniowaniem o ściśle określonej długości fali (kolejne barwy tęczy).
- **Barwy achromatyczne** (niekolorowe) – odcienie szarości od bieli do czerni

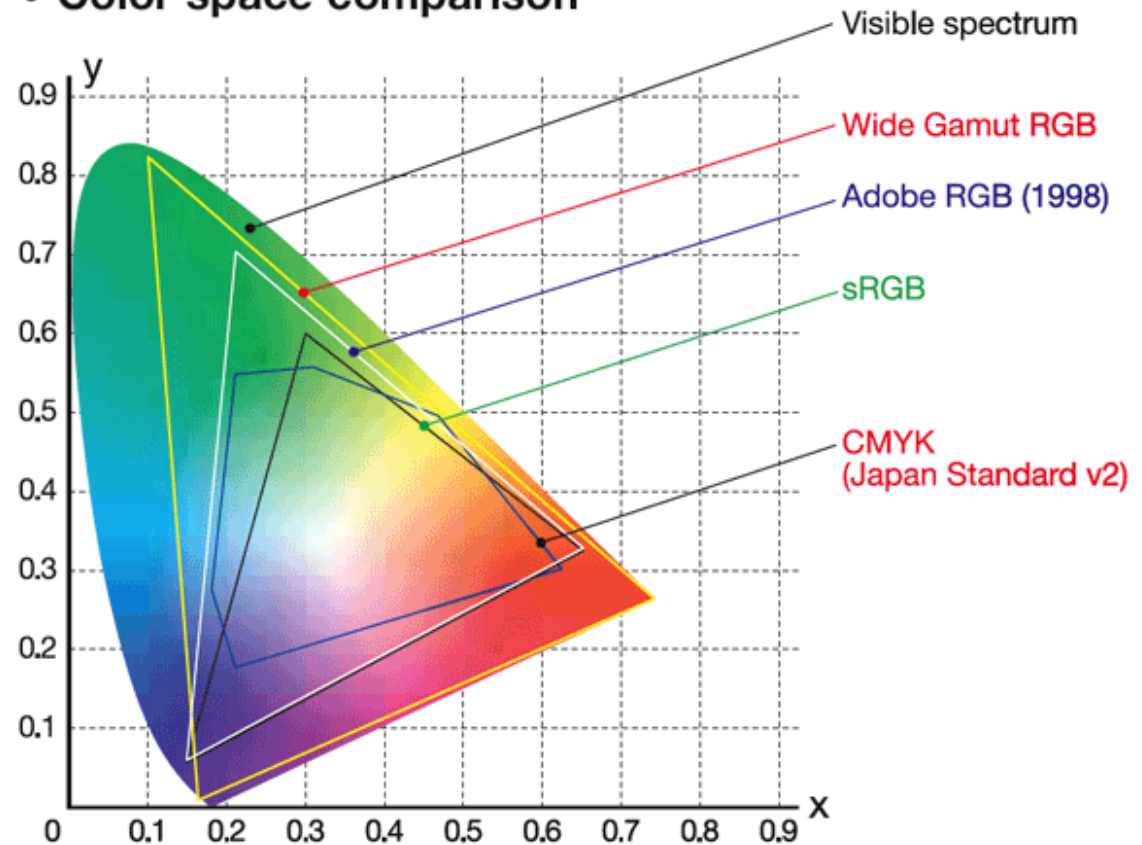
Mieszanie kolorów – 3 prawo Grassmana

- Trzy barwy tworzą **układ niezależnych liniowo barw (barwy podstawowe)**, jeżeli dowolne zsumowanie dwóch z nich nie może dać trzeciej barwy układu

- Prawo Grassmana:

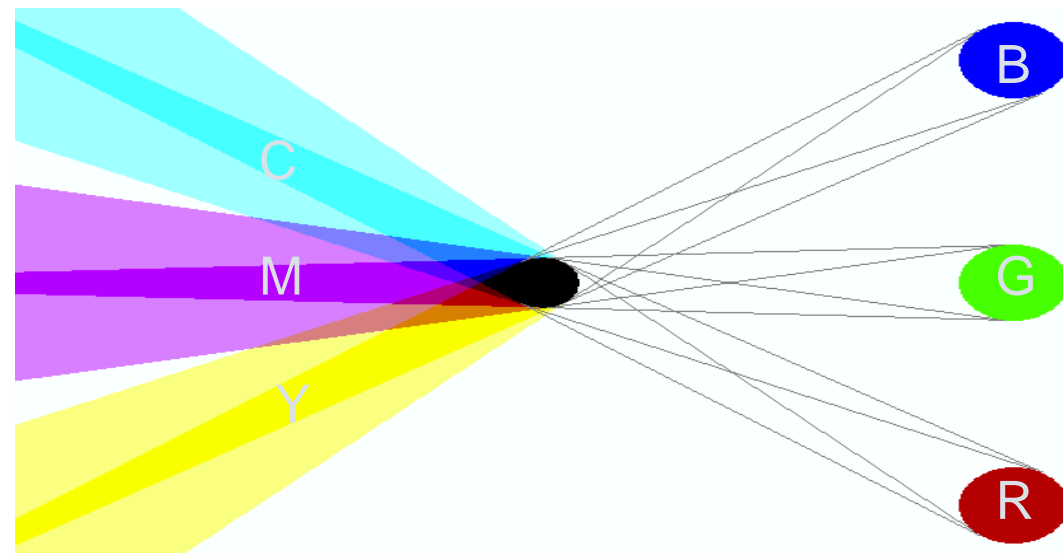
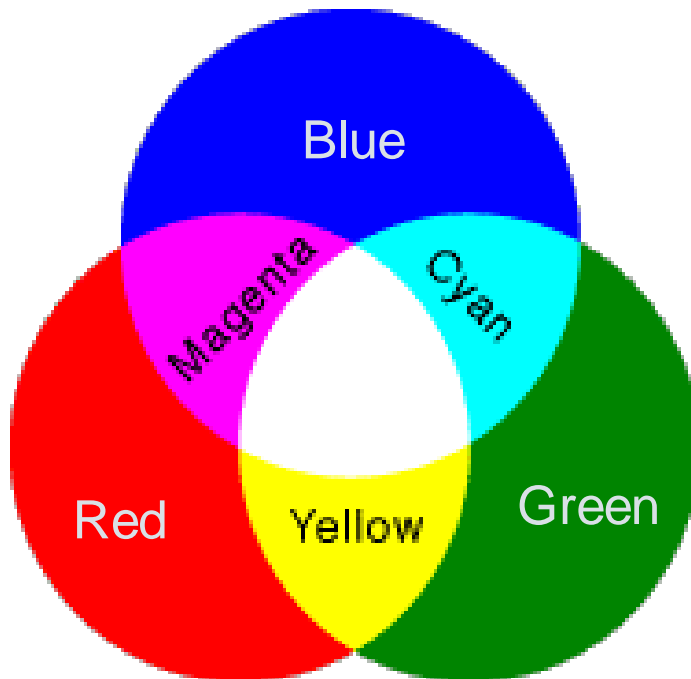
Każdą dowolnie wybraną barwę można otrzymać za pomocą trzech liniowo niezależnych barw (podstawowych)

• Color space comparison



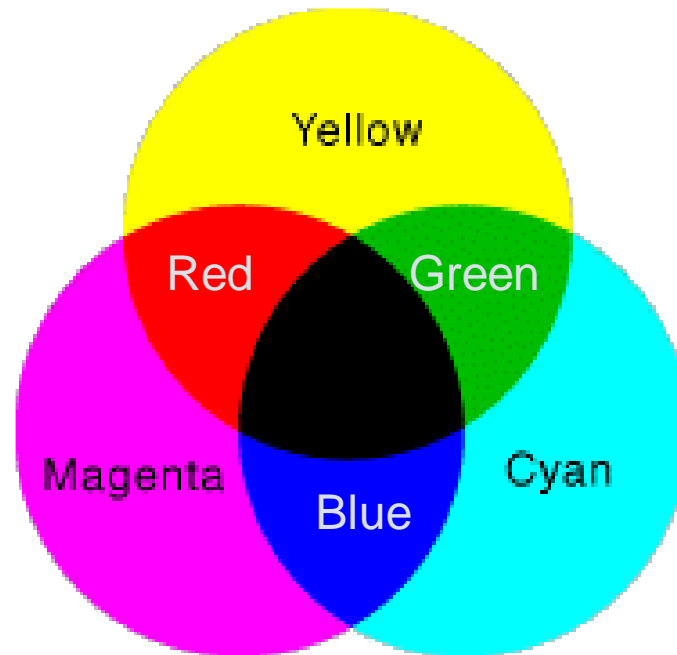
Addytywne mieszanie kolorów

- Uzyskiwanie barw poprzez mieszanie **źródeł światła**
- Światło padające bezpośrednio z kilku źródeł
- Sprzęt: CRT (monitory, TV), OLED

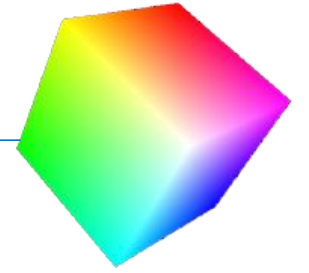


Subtraktywne uzyskiwanie kolorów

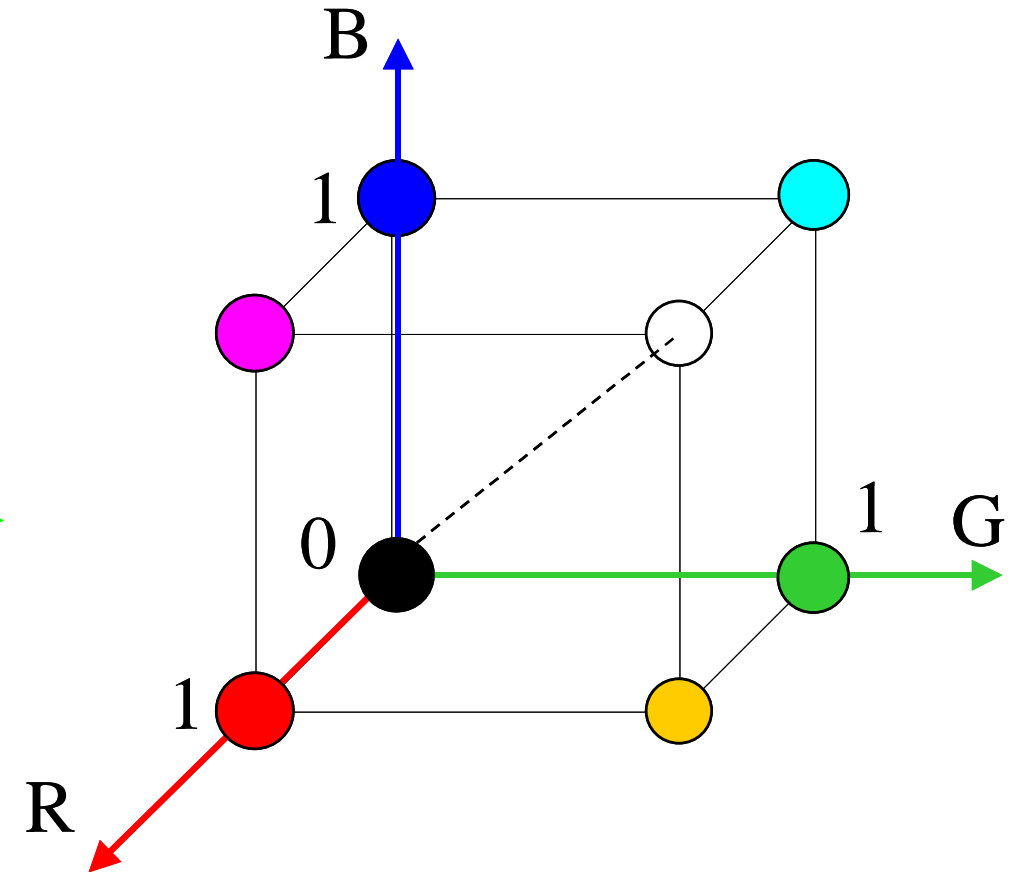
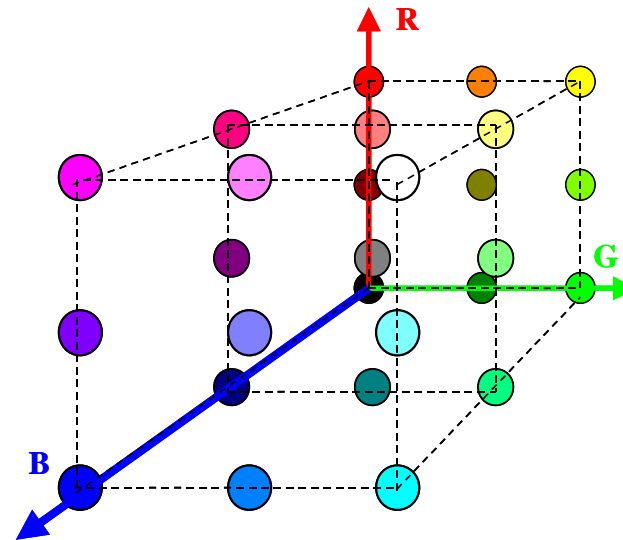
- Uzyskiwanie barw poprzez eliminację ze światła białego niektórych długości fal
- Światło odbite lub przefiltrowane
- Sprzęt: drukarki



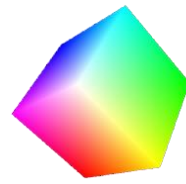
Model RGB



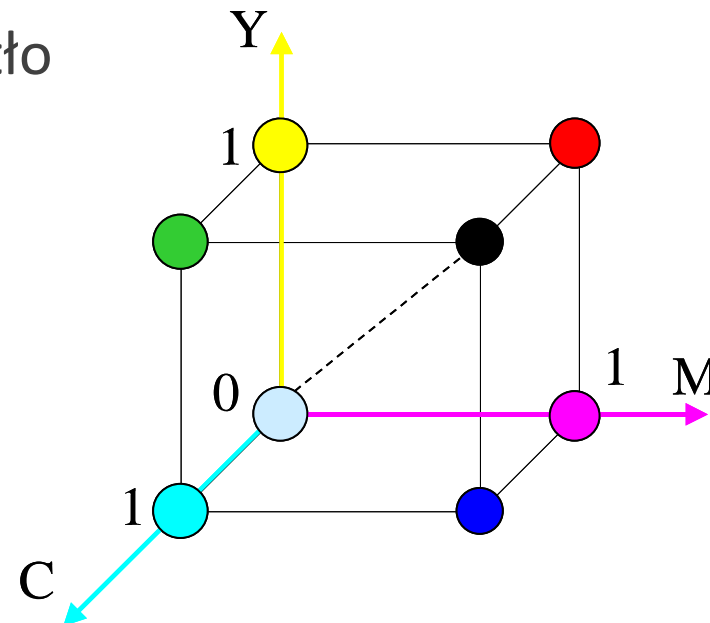
- Układ trzech, niezależnych liniowo kolorów RGB (*red, green, blue*)
- Model ukierunkowany na sprzęt
- Sprzęt emitujący światło
 - monitory
 - rzutniki



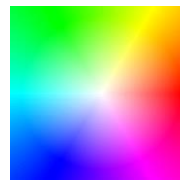
Model CMY(K)



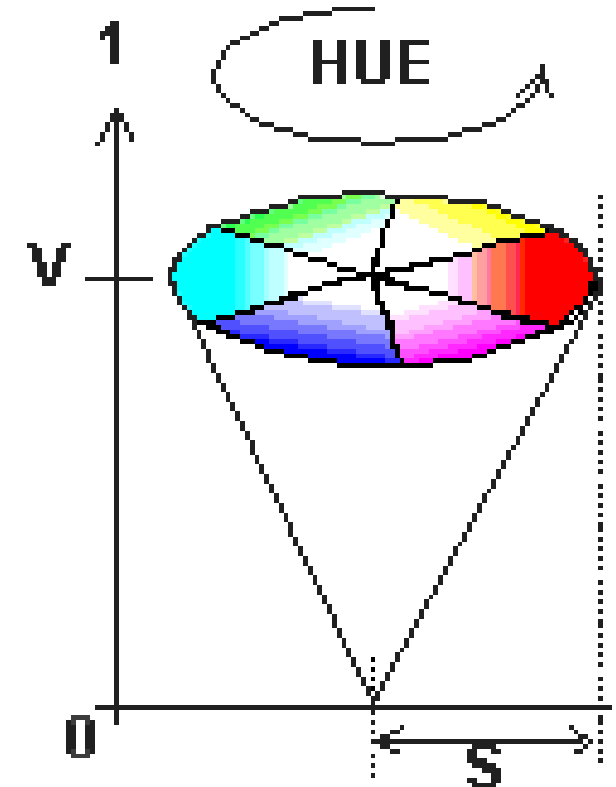
- Układ trzech, niezależnych liniowo kolorów CMY (*cyan, magenta, yellow*)
- Dodatkowy kolor K (*Key=black*), potrzebny ze względu na:
 - niedoskonałość materiałów
 - koszty
- Model ukierunkowany na sprzęt
- Sprzęt tworzący materiały odbijające światło
 - drukarki, plotery, maszyny offsetowe



Model HSV



- Układ trzech parametrów HSV
 - H (*Hue*) – częstotliwość fali światła
 - S (*Saturation*) – nasycenie barw
 - V (*Value = Brightness = Intensity*) – poziom światła białego
- Model ukierunkowany na użytkownika (artysta)
- Układ współrzędnych (odwrócony stożek lub ostrosłup)



1.1.2. Dźwięk

- **Dźwięk** - fala akustyczna rozchodząca się w ośrodku sprężystym powodująca zmianę ciśnień (ruch cząsteczek) ośrodka
- **Fale akustyczne** (sprężyste) - zaburzenia mechaniczne (fala) ośrodka sprężystego, nie powodujące przesunięcia średnich położenia atomów ośrodka
- W cieczech i gazach fala akustyczna jest falą podłużną, w ciałach stałych może być zarówno falą podłużną, jak i poprzeczną.

Rodzaje dźwięków

- Posiadające określoną wysokość
 - **ton** (dźwięk prosty) - harmoniczne (sinusoidalne) drganie o jednej częstotliwości
 - **wieloton harmoniczny** – złożony z tonu podstawowego i częstotliwości harmoniczných (wielokrotności tonu podstawowego)
- Nieposiadające określonej wysokości
 - **wieloton nieharmoniczny** – złożony ze składowych nieharmonicznych
 - **szmer** – odgłosy o nieokreślonej wysokości
 - **szum** - dźwięk o ciągłym widmie

Charakterystyka dźwięku

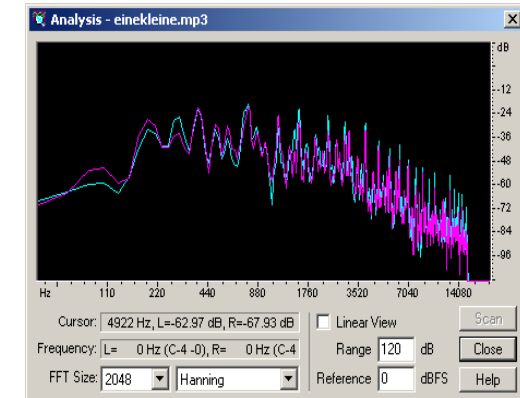
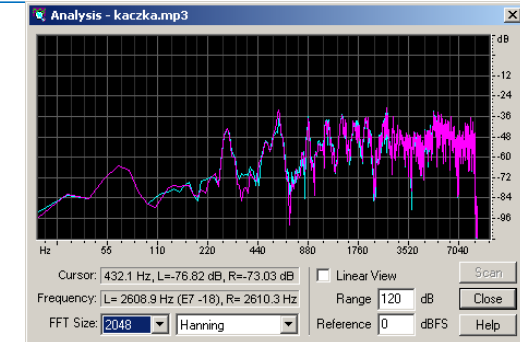
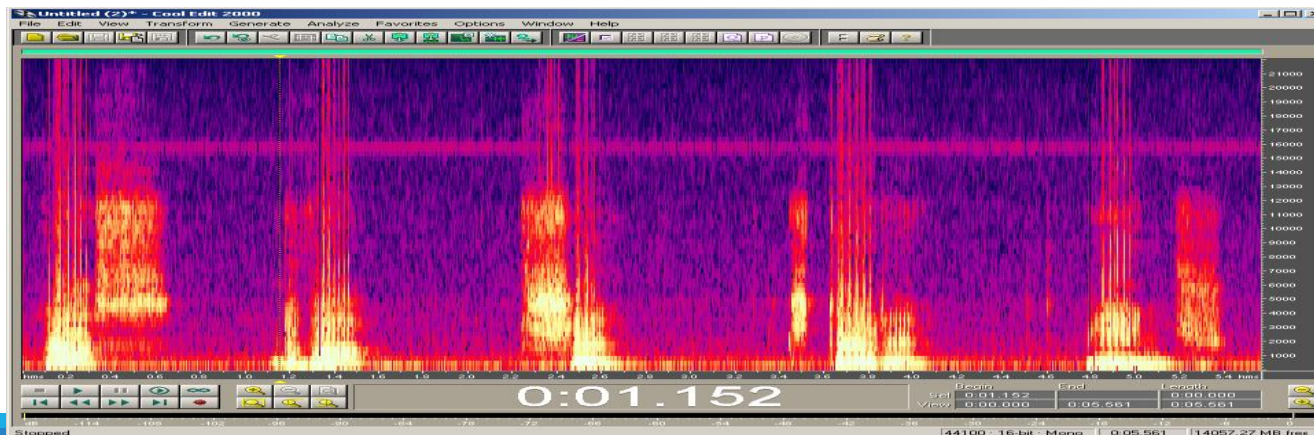
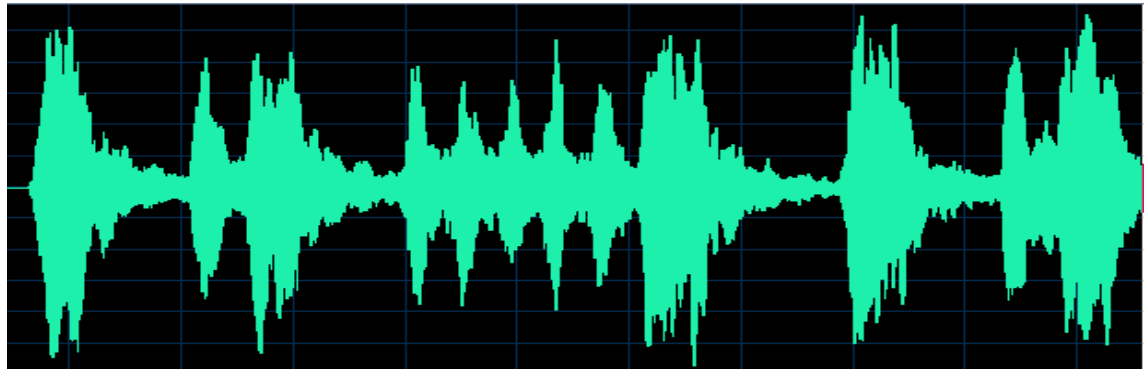
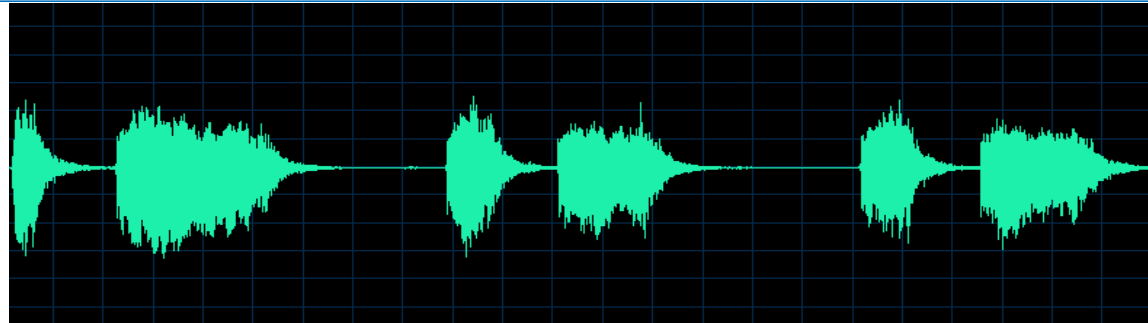
- **Wysokość** - częstotliwość tonu podstawowego
- **Natężenie I** - ilość energii przenoszona w jednostce czasu przez prostopadłą jednostkę powierzchni [W/m²]

$$I = \frac{p^2}{2\rho v}$$

, gdzie: p - amplituda ciśnienia akustycznego, ρ - gęstość ośrodka, v - prędkość rozchodzenia się fali akustycznej

- **Barwa** - proporcje natężeń harmoniczných

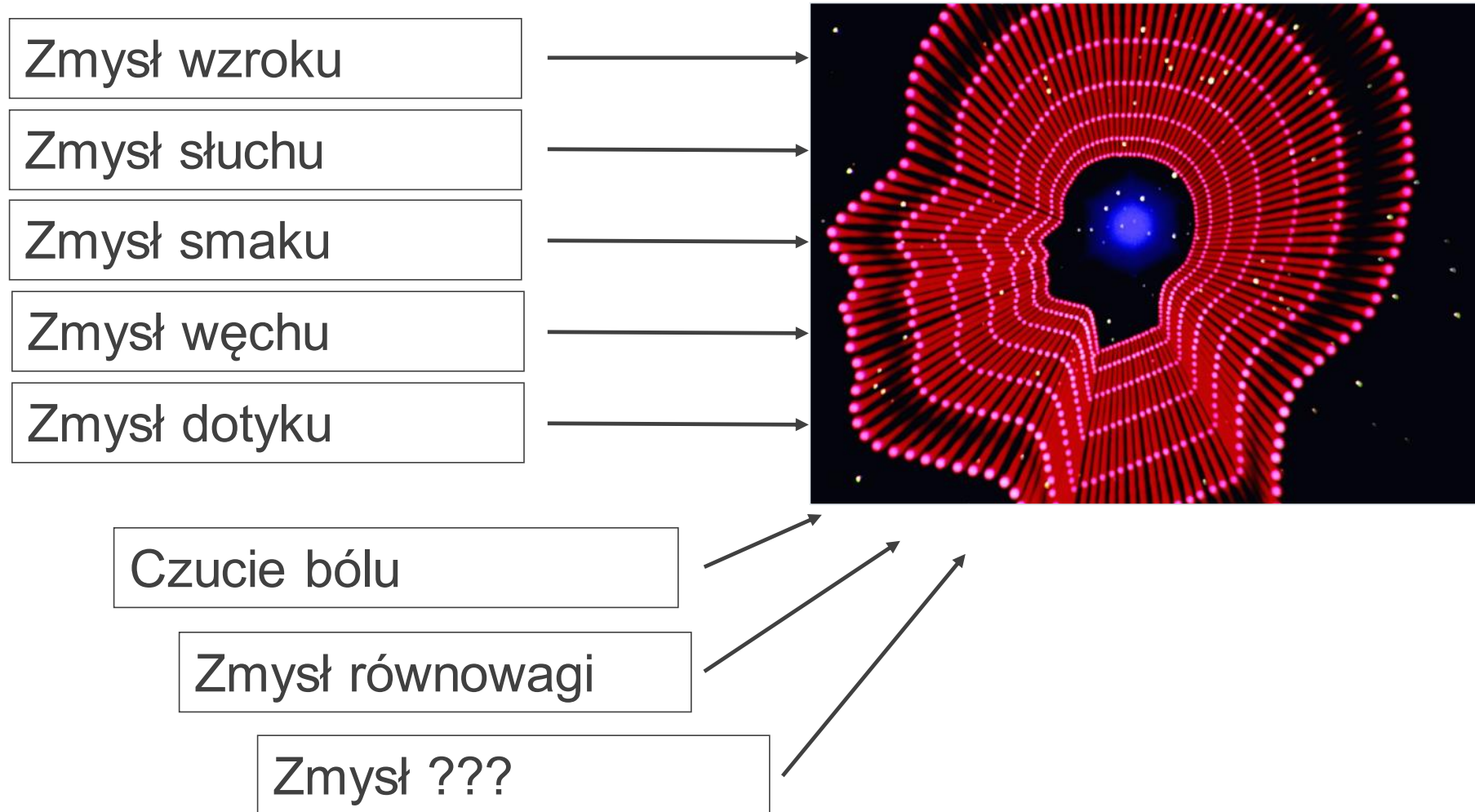
Przebieg i widmo sygnału mowy



Rozchodzenie się fal akustycznych

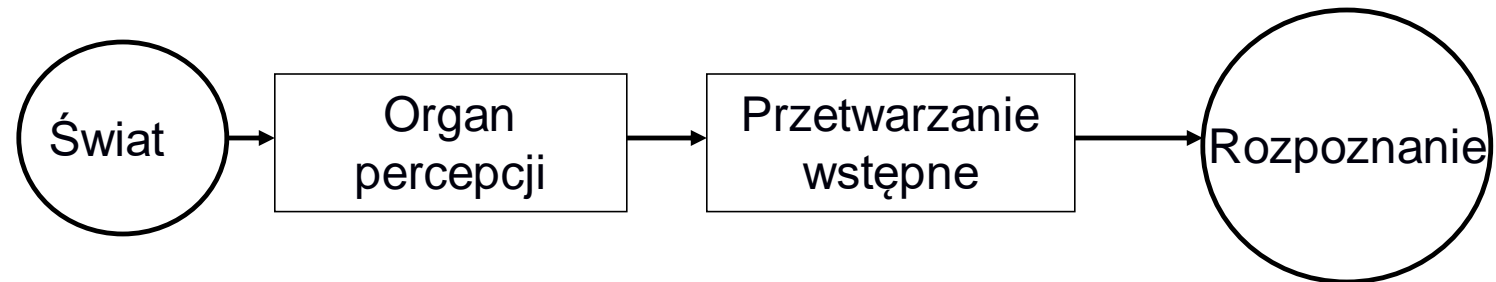
- **Odbicie** - od przeszkody
- **Echo** - powrót odbitej fali do źródła
- **Załamanie** przy przejście do ośrodka o innej prędkości rozchodzenia się fal
- **Dyfrakcja** - ugięcie fali w pobliżu przeszkody
- **Interferencja** - nakładanie się fal o identycznej częstotliwości
- **Dudnienia** - nakładanie się fal o nieznacznie różniących się częstotliwościami
- **Fala stojąca** - specyficzna interferencja fal o identycznych parametrach, lecz przeciwnych kierunkach (np. padającej i odbitej)

1.2. Podstawy percepcji człowieka



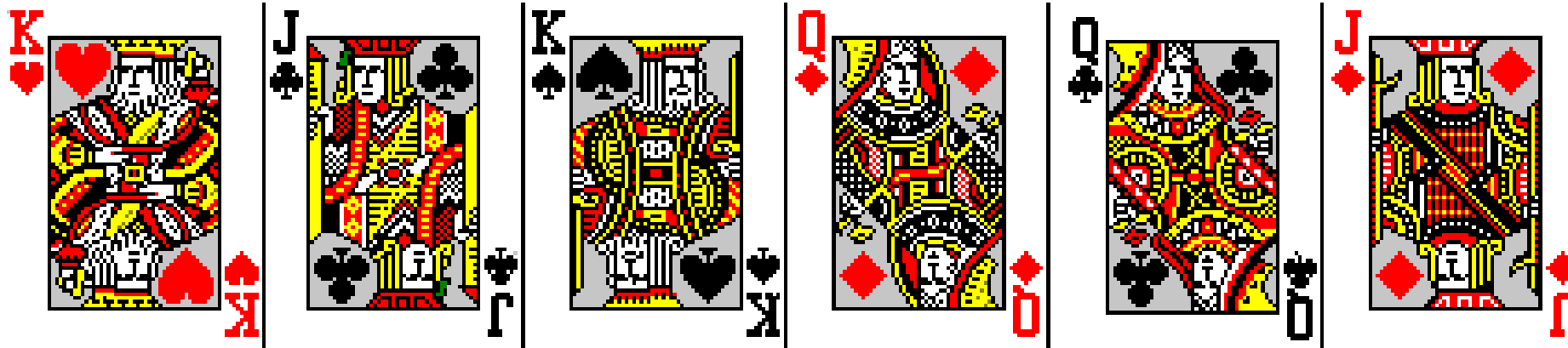
Mózg jako organ percepcji

- Zadania mózgu
 - rejestracja i zapamiętywanie danych/informacji
 - filtrowanie (maskowanie) danych oraz informacji
 - wydobywanie informacji
 - uzupełnienie brakujących danych/informacji
 - kojarzenie
 - koordynacja
 - reakcja



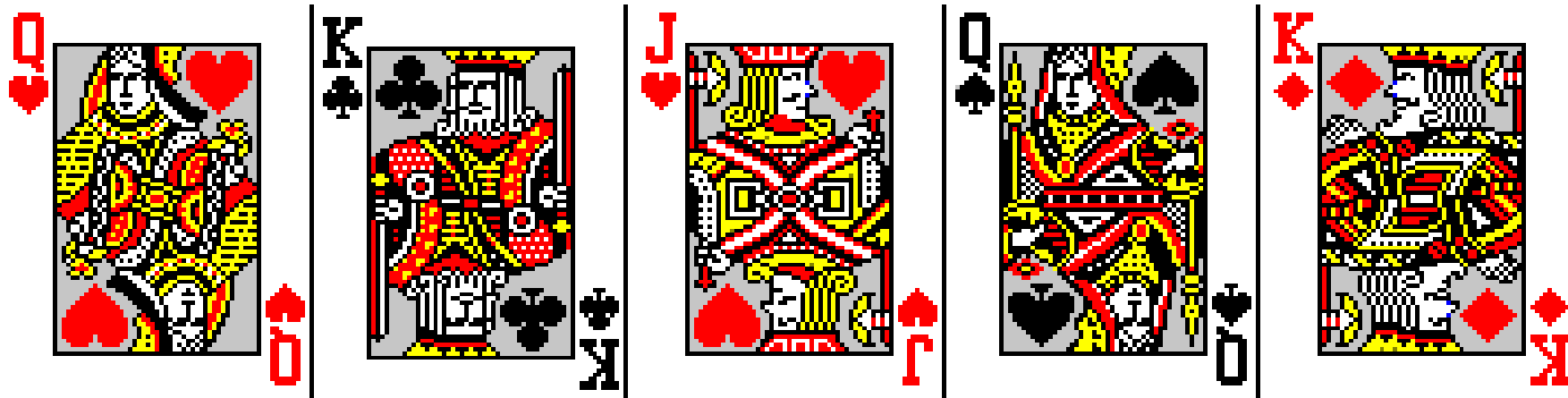
Informacja nadmiarowa (1)

Proszę wybrać dowolną kartę

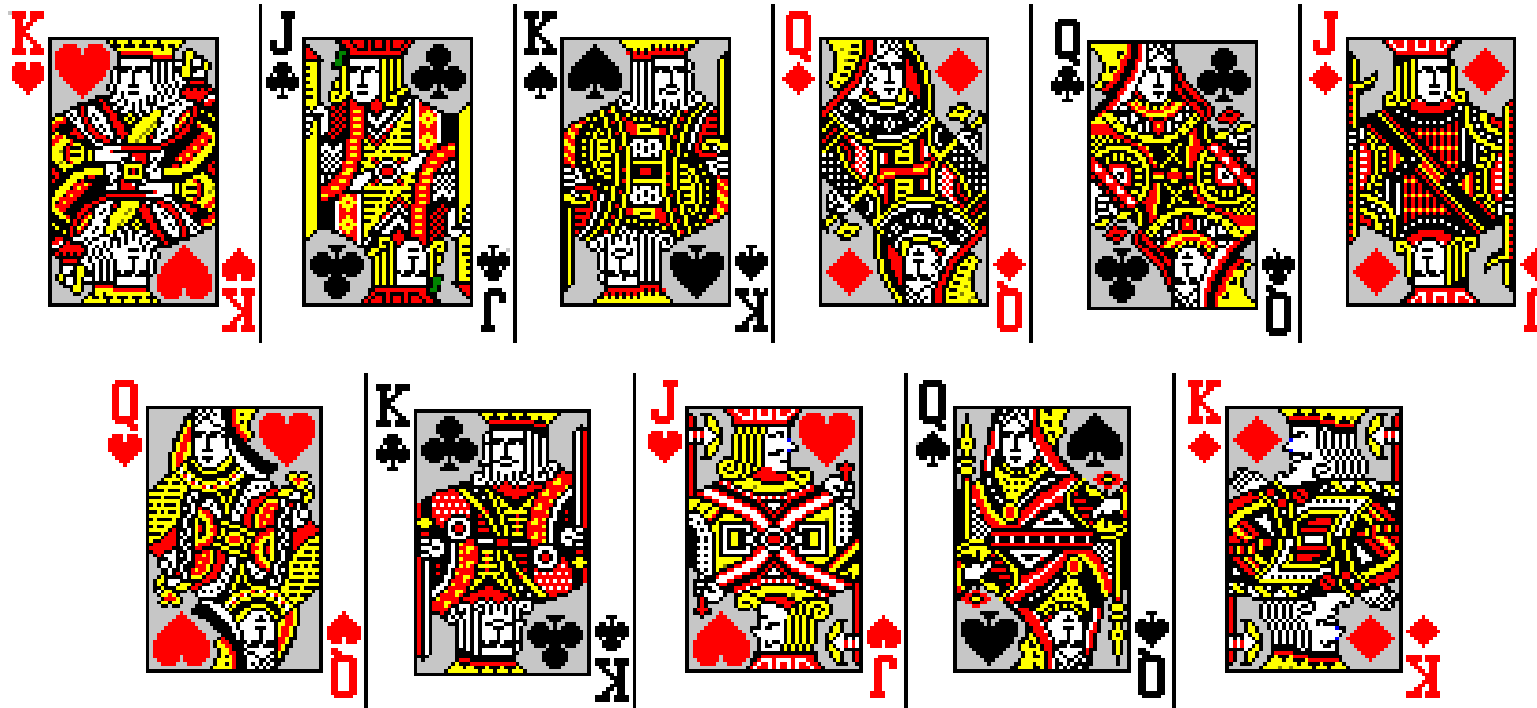


Informacja nadmiarowa (2)

- Wybrana karta została usunięta



Wyjaśnienie



- Mózg nie radzi sobie za dobrze z całym strumieniem informacji przychodzącej, dlatego wrażenia optyczne są w nim filtrowane i tylko nieznaczna część informacji zawartej w otoczeniu jest rozumiana i interpretowana

Percepcja inaczej

- Idealizm subiektywny (solipsyzm): *istnieję tylko ja i moja świadomość; świat może istnieć jako wyobrażenie lub sen*
 - *filozofia, literatura i film s-f*
- Idealizm obiektywny: *świat jest jedynie odbiciem uniwersalnej idei - bytu duchowego*
 - *filozofia, sztuka, malarstwo*

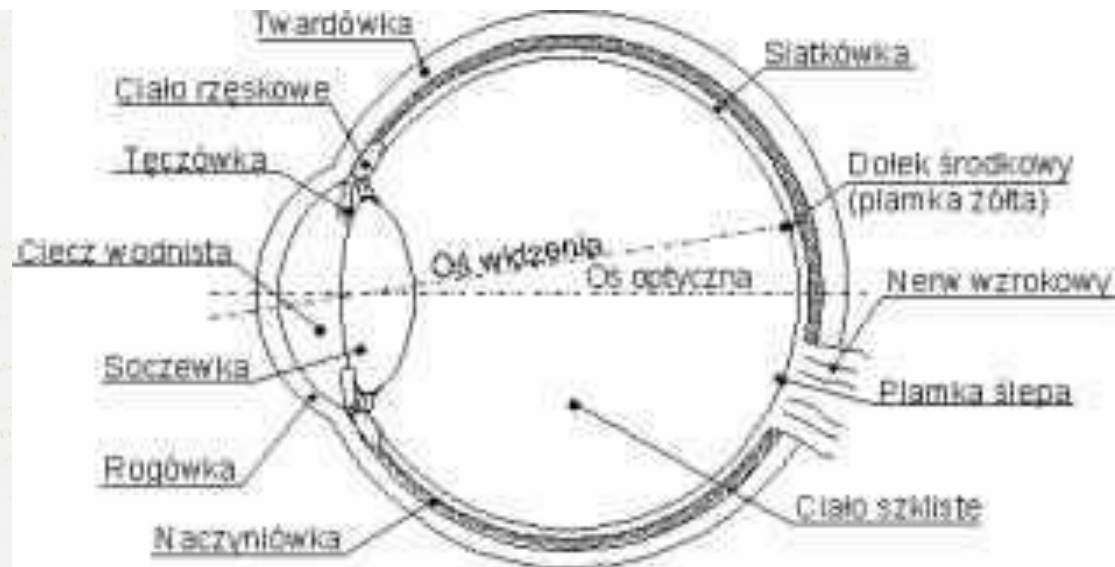
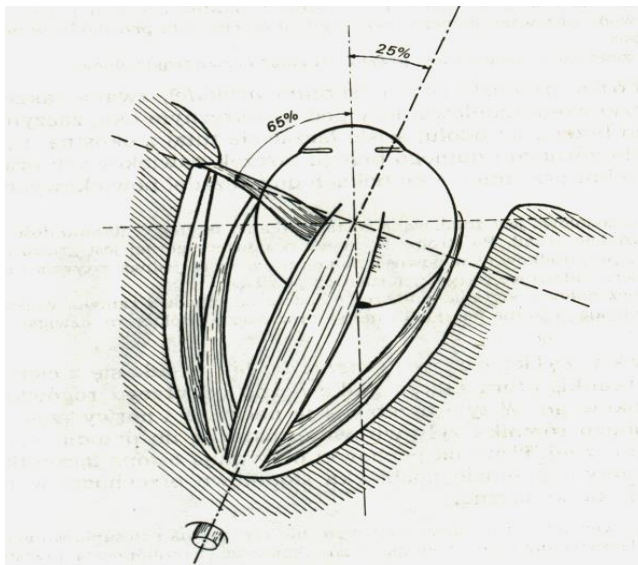
1.2.1. Zmysł wzroku

- **Wzrok** jest zmysłem pozwalającym na percepcję promieniowania elektromagnetycznego w zakresie widzialnym (400-700 nm)
- Wzrok reaguje na: jasność, barwę, nasycenie i ruch



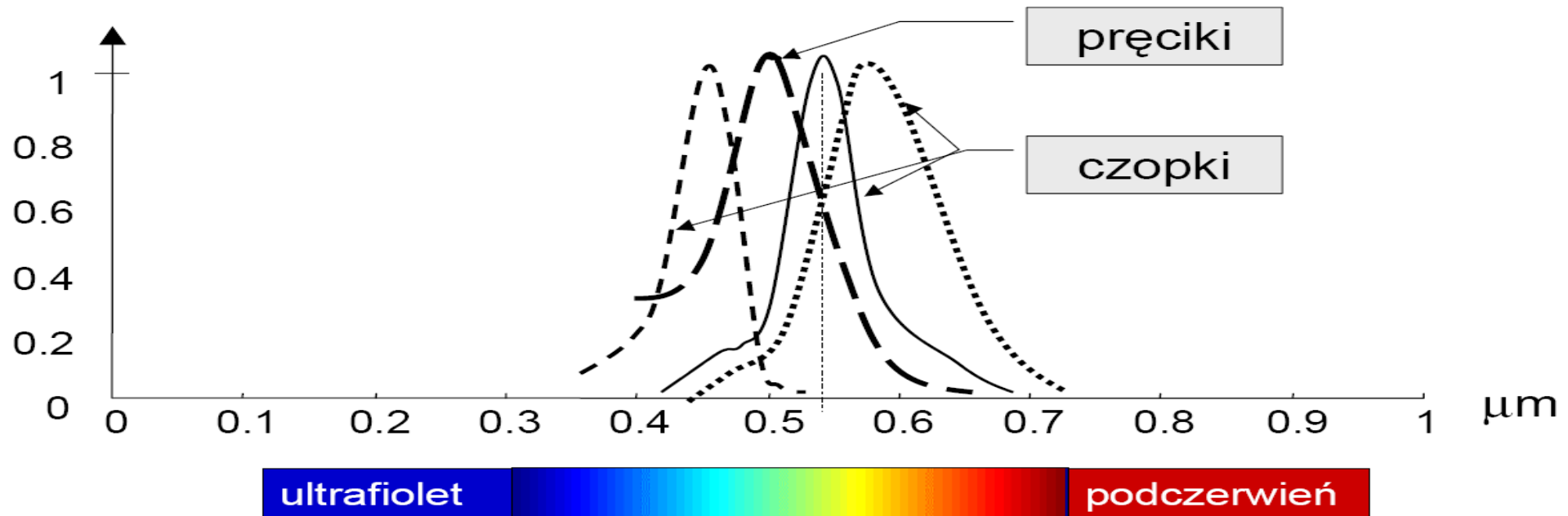
Budowa oka

- Mechanizm soczewki
- Komórki światłoczułe (receptory)
 - pręciki (odbiór jasności)
 - czopki (odbiór koloru)
- Nerw wzrokowy



Budowa oka - pręciki

- Odpowiedzialne za odbiór jasności (szeroki zakres czułości spektralnej)
- Wysoka czułość (od ok. 0.01 lux)
- Widzenie nocne (skotopowe)
- Percepcja kształtów, zarysów



Budowa oka - czopki

- Odpowiedzialne za odbiór koloru (wąski zakres czułości spektralnej)
- Widzenie koloru przy dobrym oświetleniu (od ok. 35 lux)
- Widzenie dzienne (fotopowe)
- Teorie percepcji barw
 - percepcja 3 barw RGB (Young-Helmholtz)
 - percepcja 3 par barw: *Black/White, Red/Green, Blue/Yellow*, lepiej wyjaśnia:
 - daltonizm (8% mężczyzn, 0.5% kobiet)
 - słabe widzenie niektórych mieszanek barw

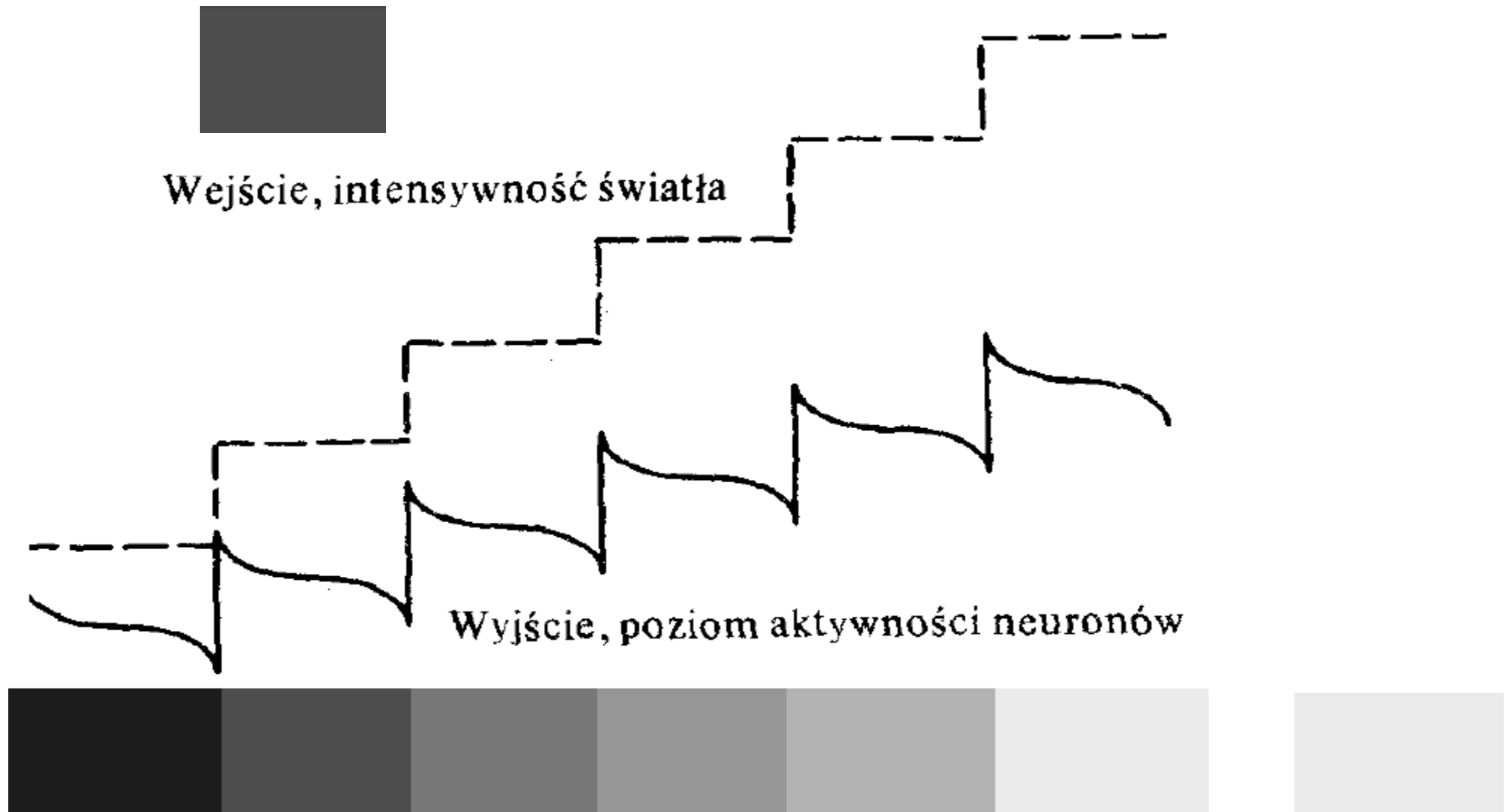
Złudzenia optyczne

Złudzenie optyczne - fałszywa interpretacja (postrzeganie) obrazu oglądanego przez człowieka spowodowana

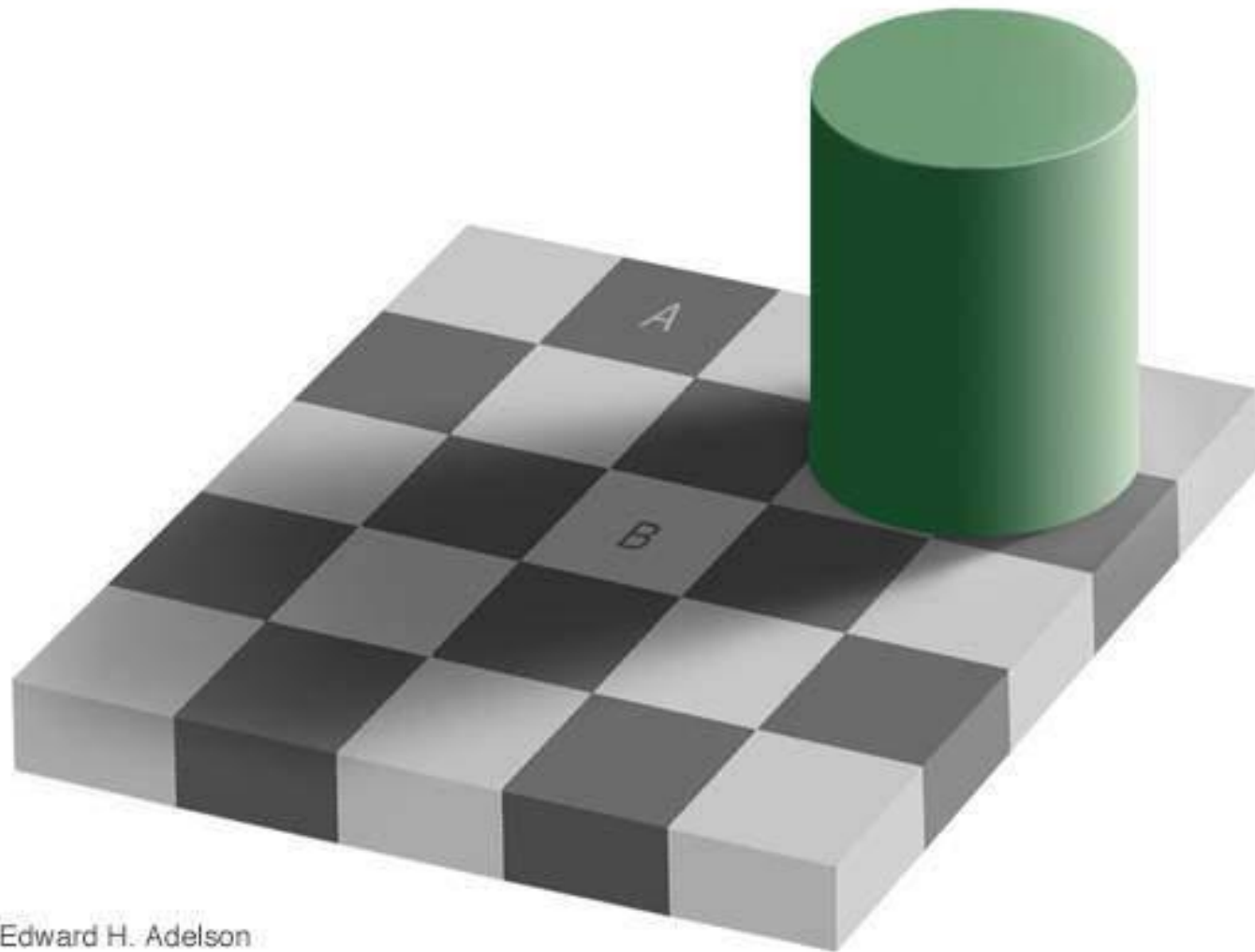
- ograniczeniami budowy oczu
- właściwościami procesu postrzegania
- kontekstowym rozpoznawaniem obrazu
- przyzwyczajeniami
- wyobraźnią
- perspektywą lub jej brakiem, ...

Złudzenie może dotyczyć takich cech jak jasność, barwa, rozmiar, odległość, prostoliniowość, równoległość i in.

Nieliniowa charakterystyka percepcji jasności



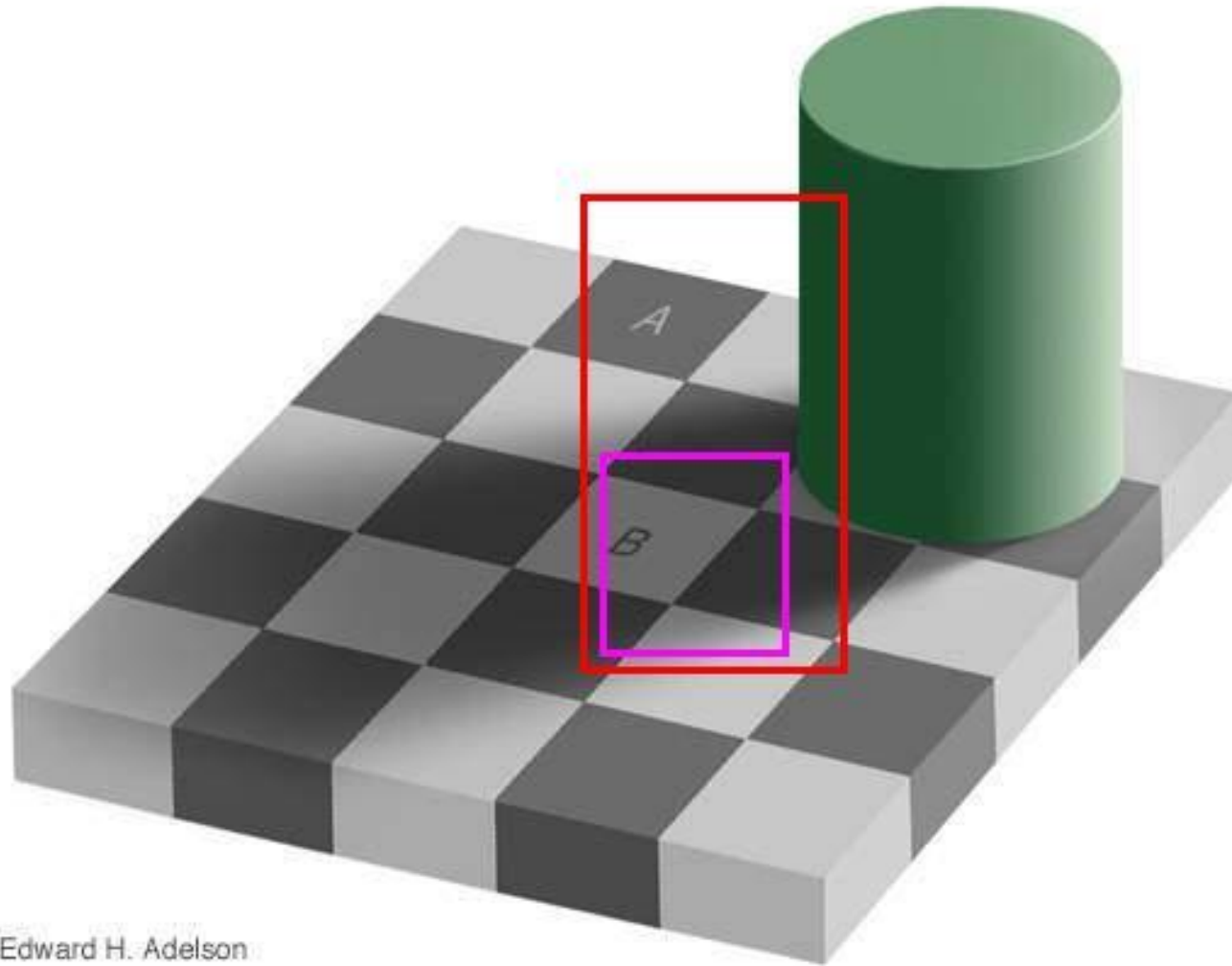
Szachownica (1)



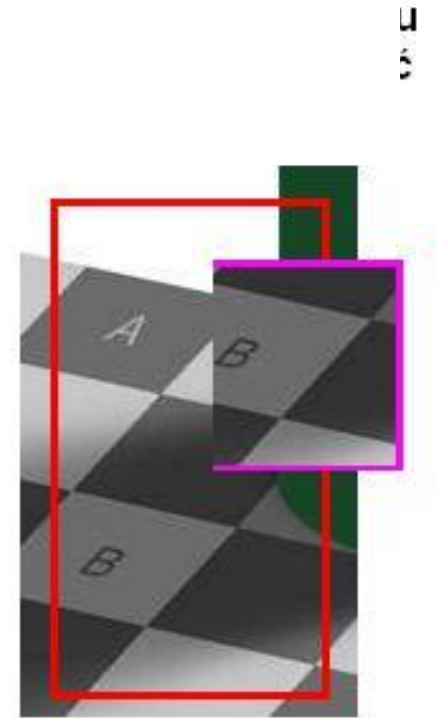
Edward H. Adelson

Szachownica (2)

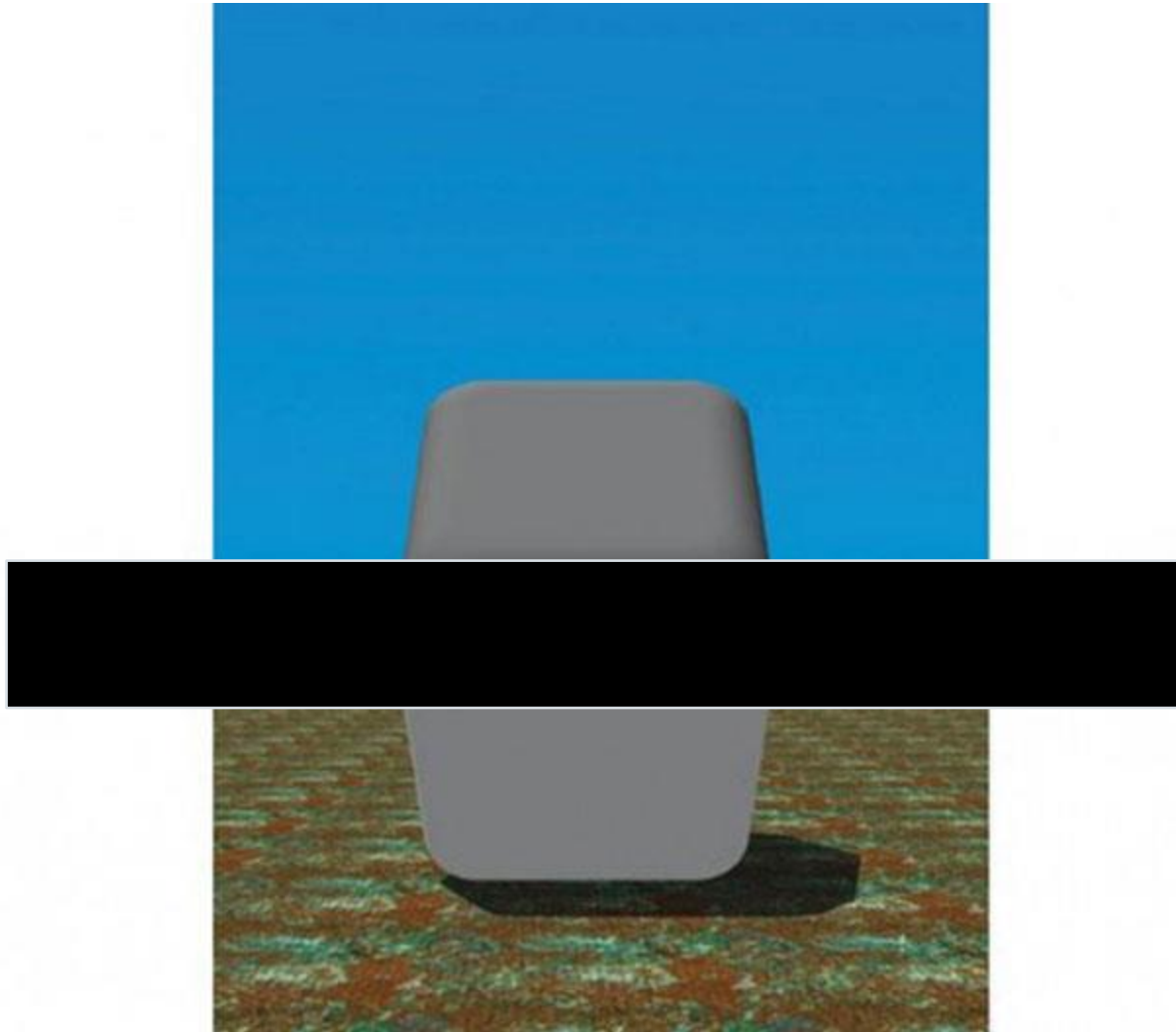
WZROK 2021 104 501 1000 1000



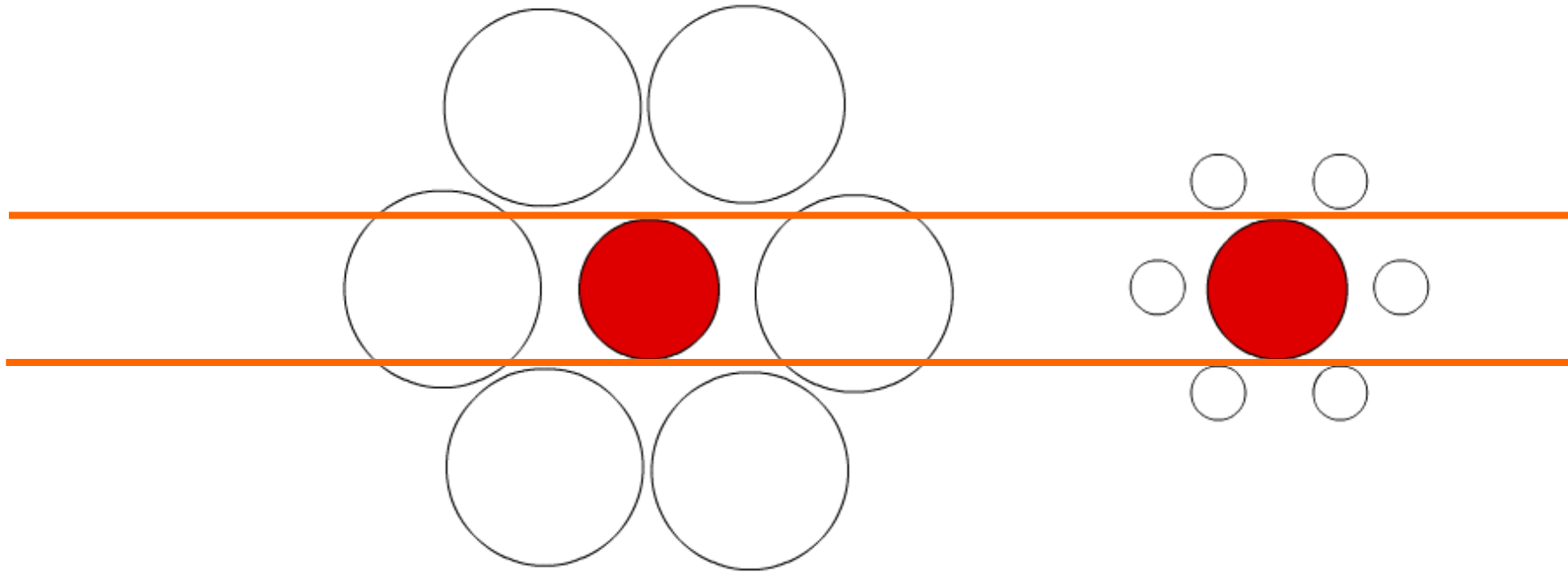
Edward H. Adelson



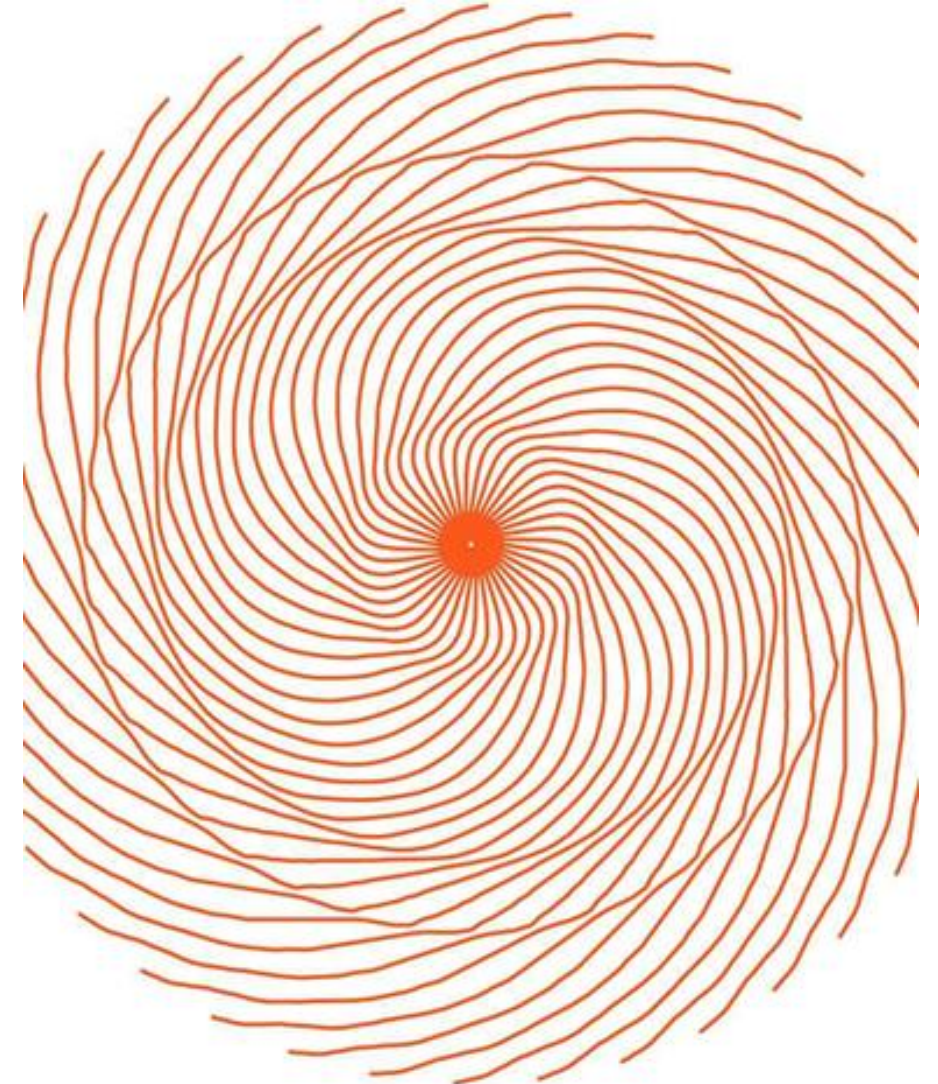
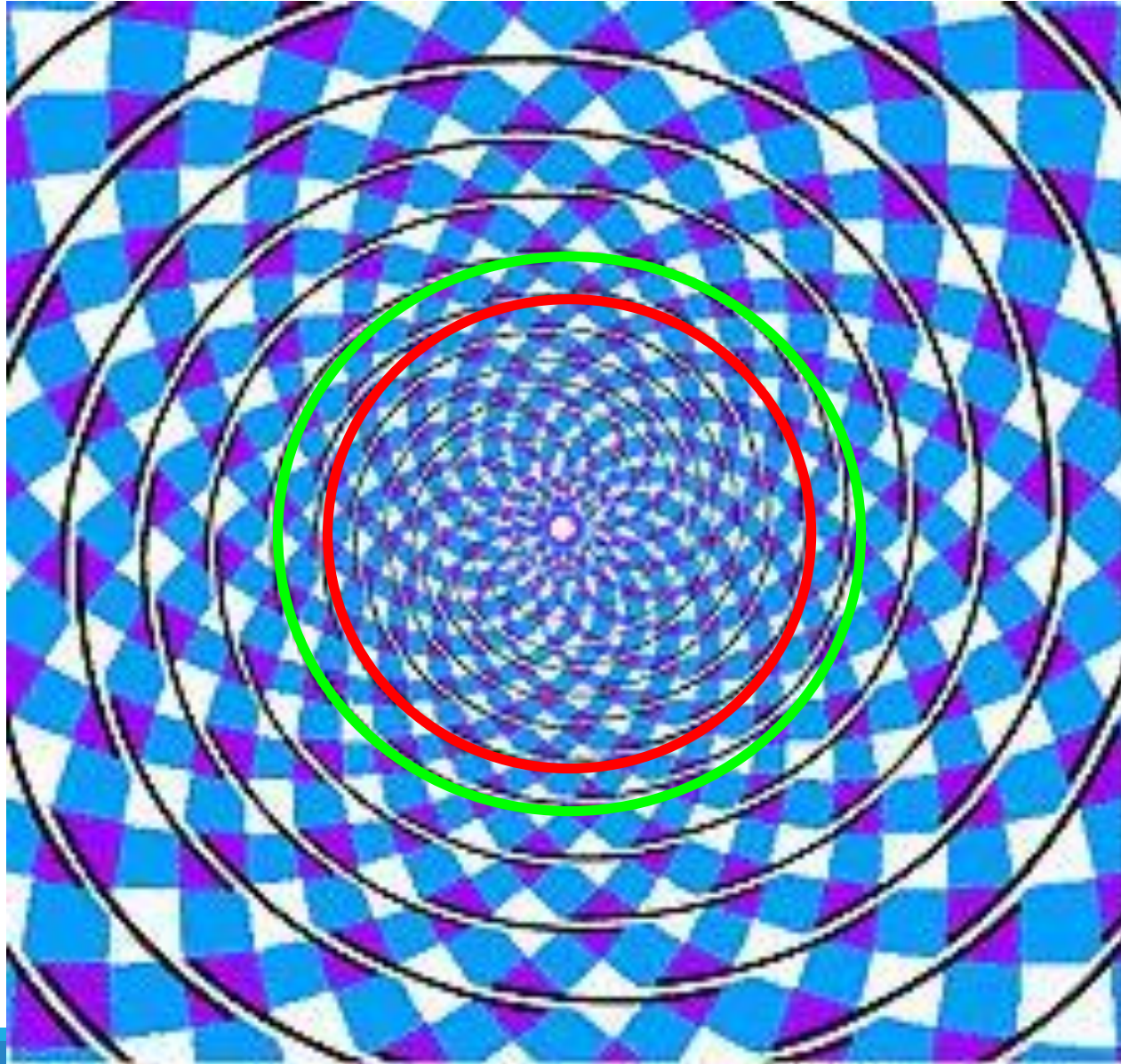
Sześćcian



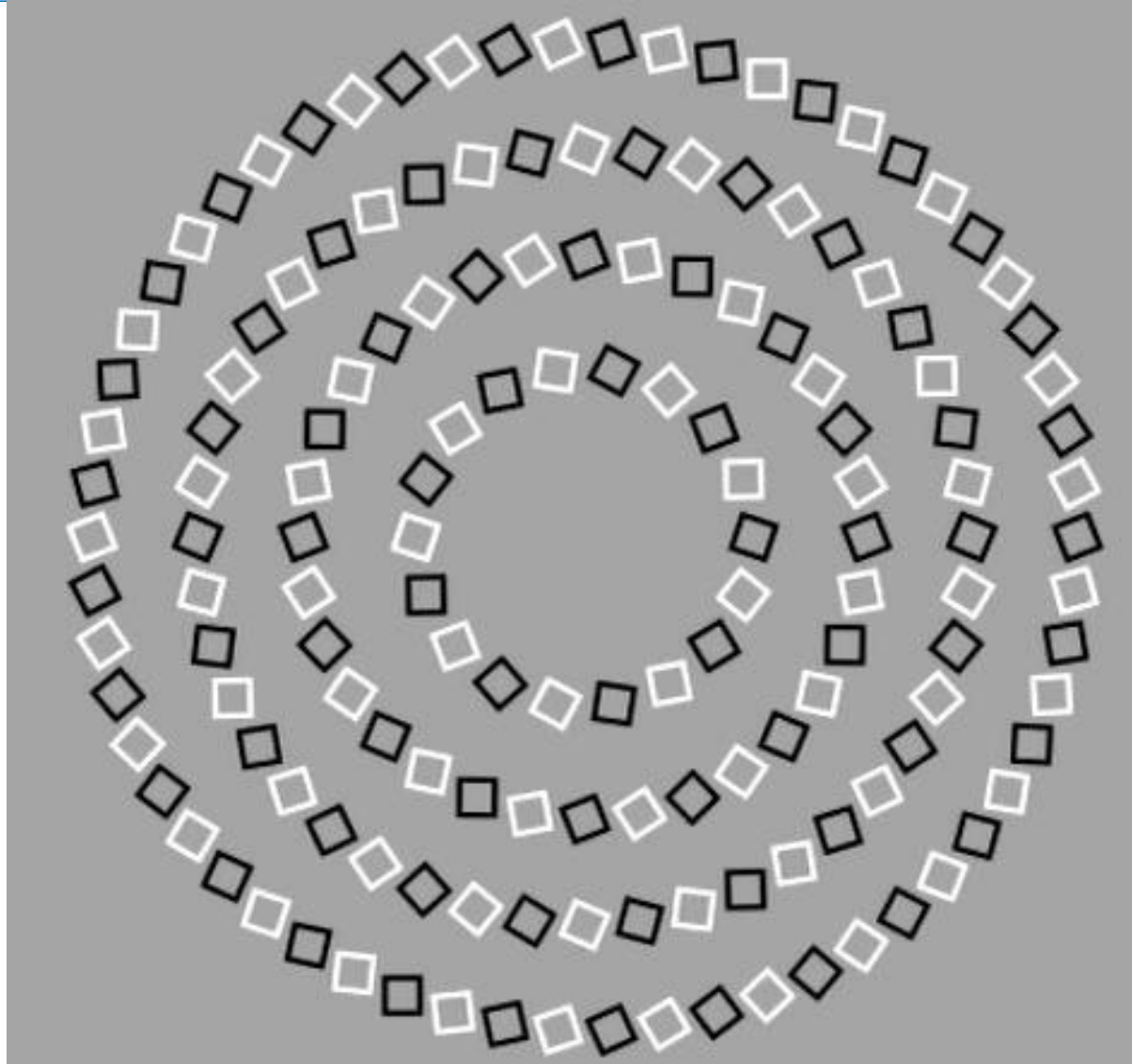
Złudzenia optyczne (3)



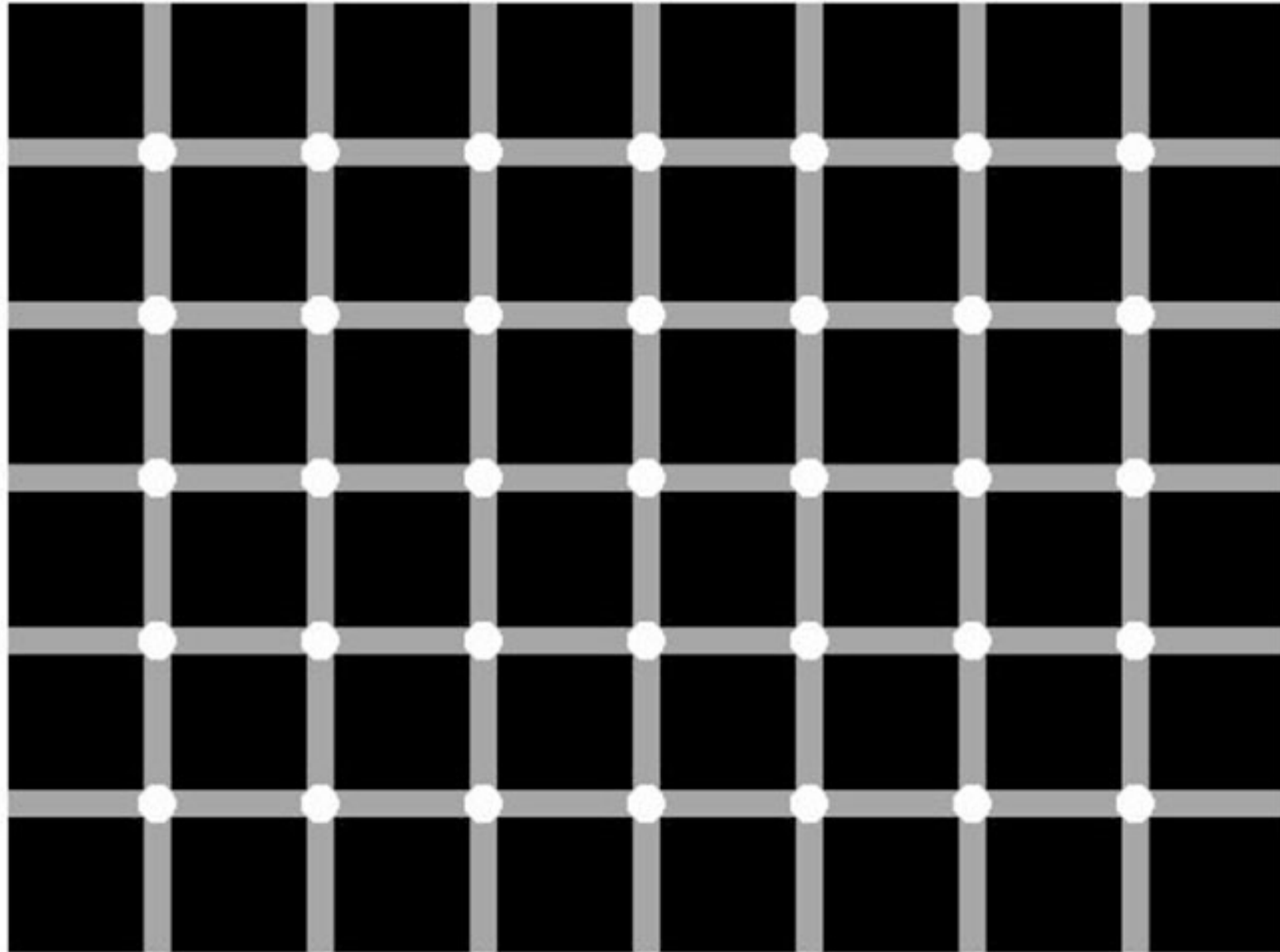
Złudzenia optyczne (4)



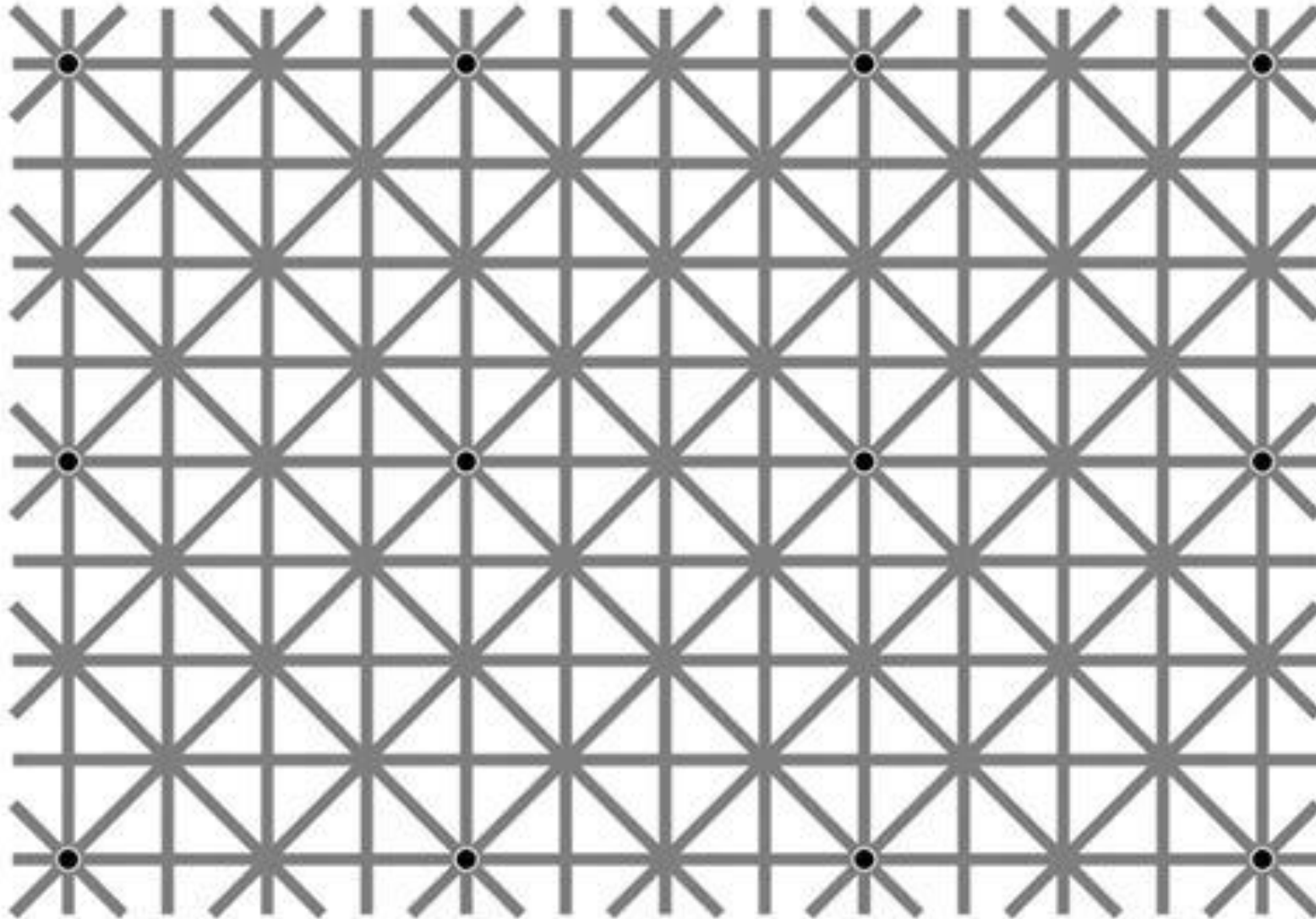
Policz okręgi ;-)



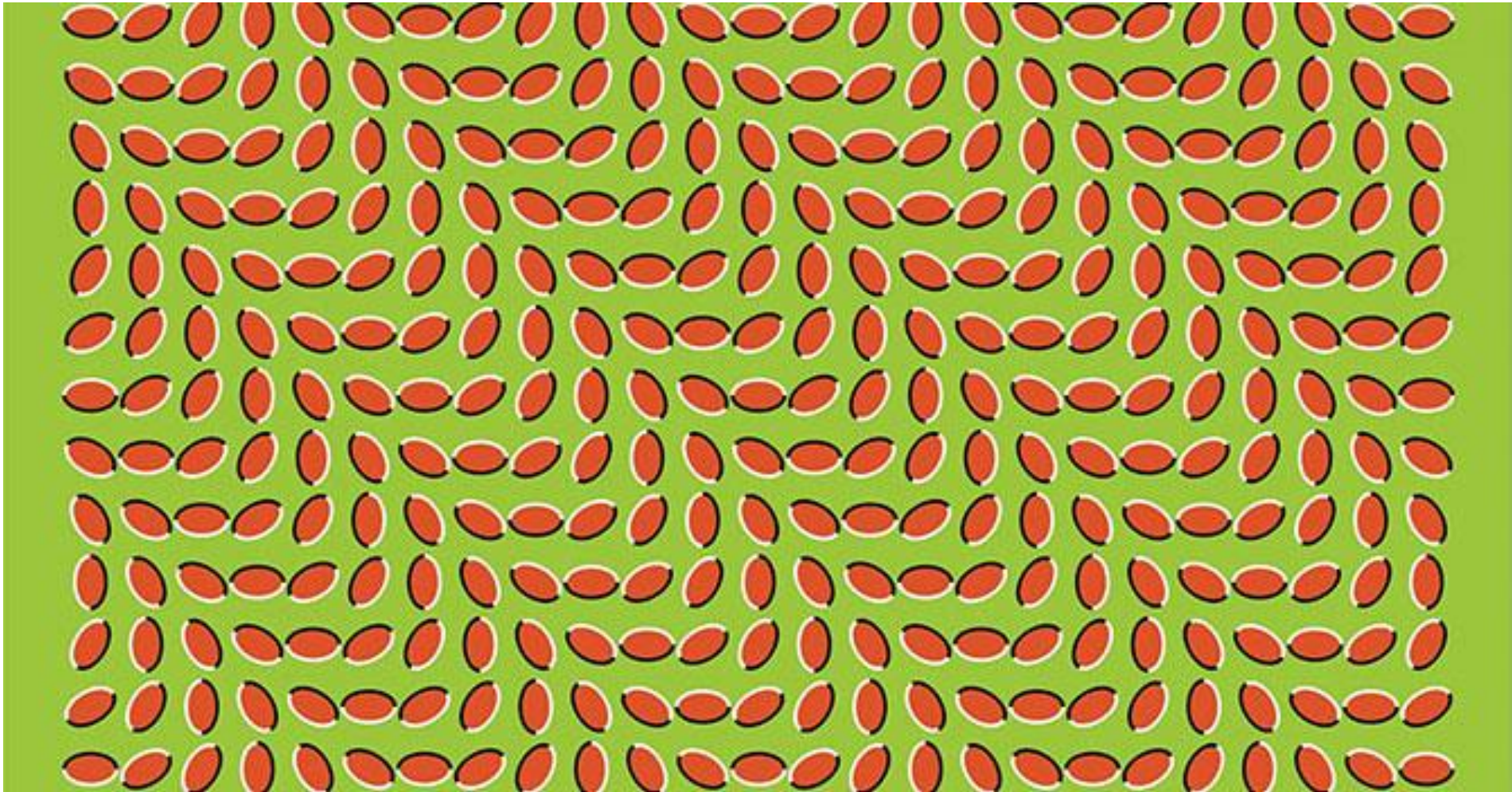
Siatka Hermanna



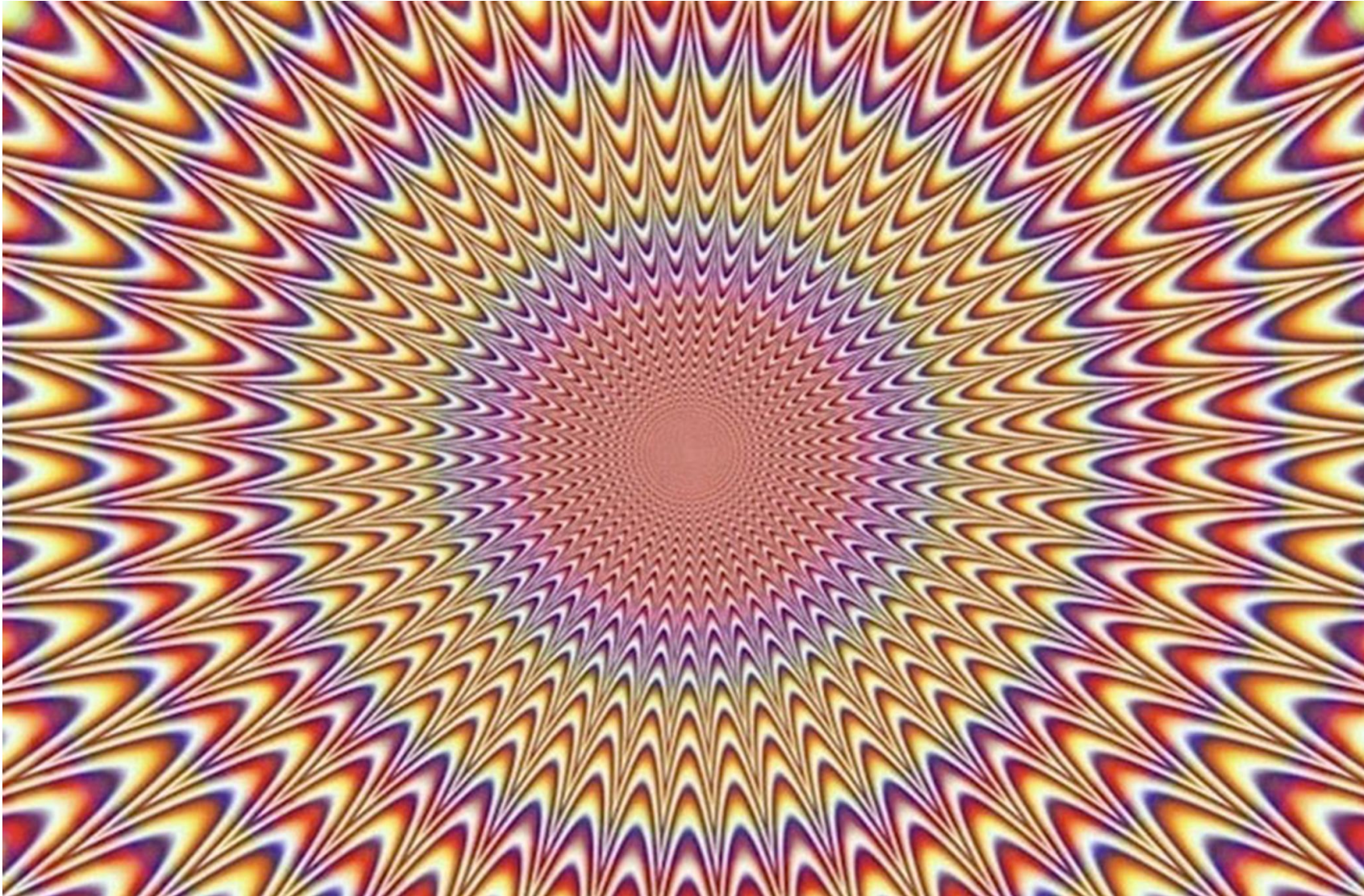
Znajdź kropki



Złudzenie ruchu



Złudzenie ruchu



Złudzenie ruchu

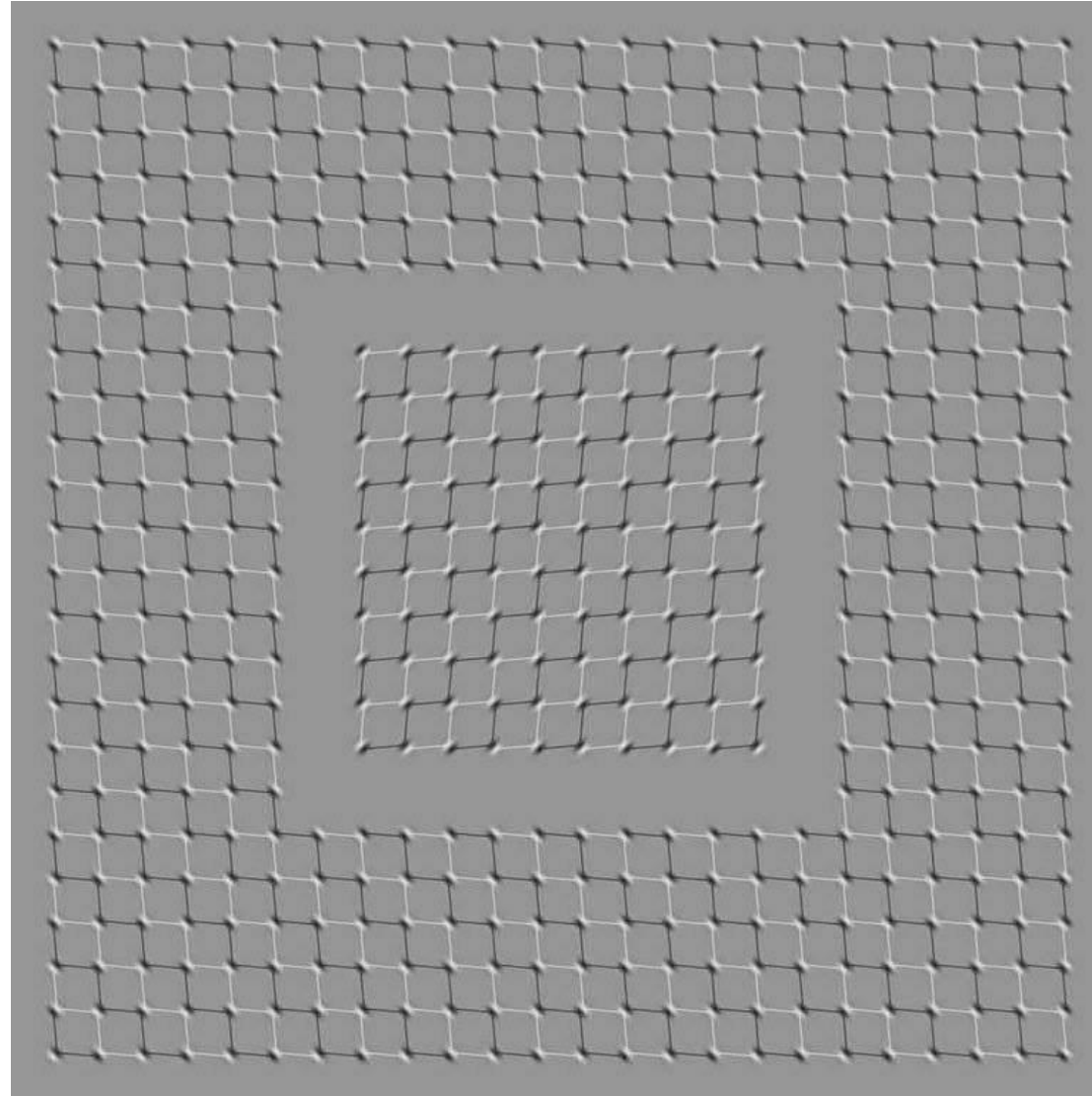


Figura Zollnera

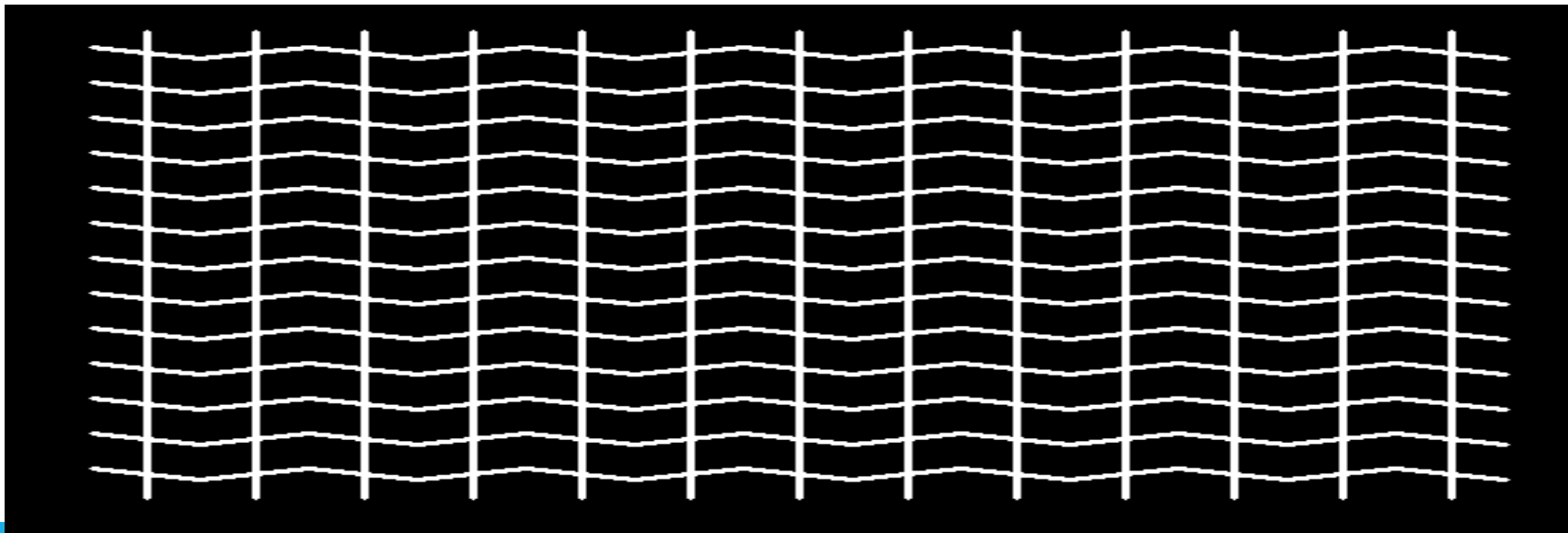
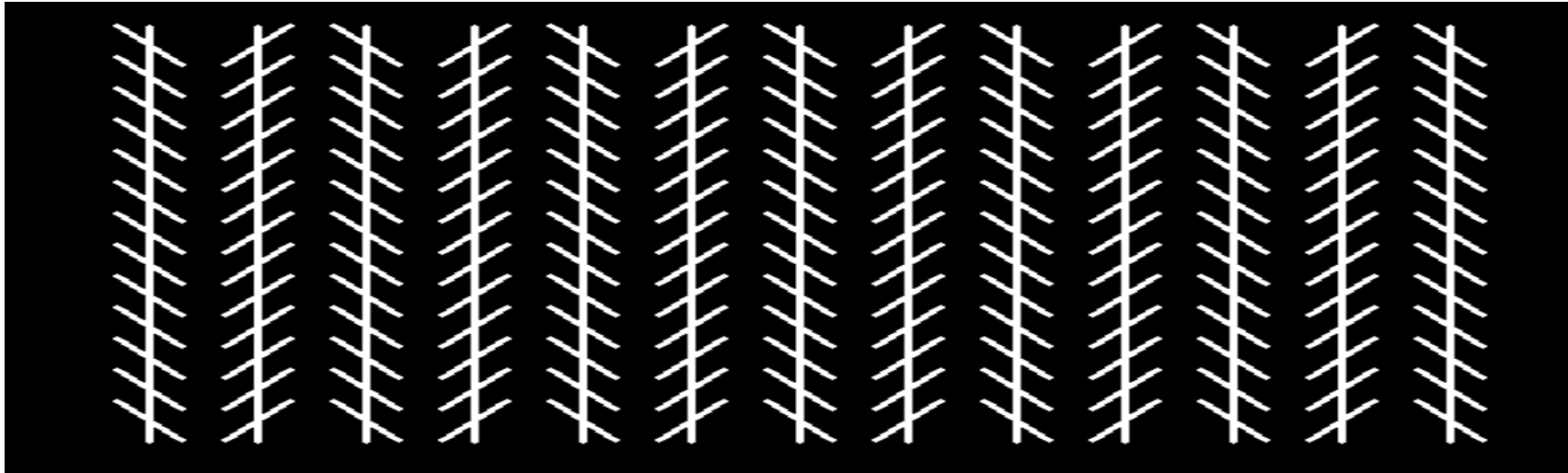


Figura Müllera-Lyera

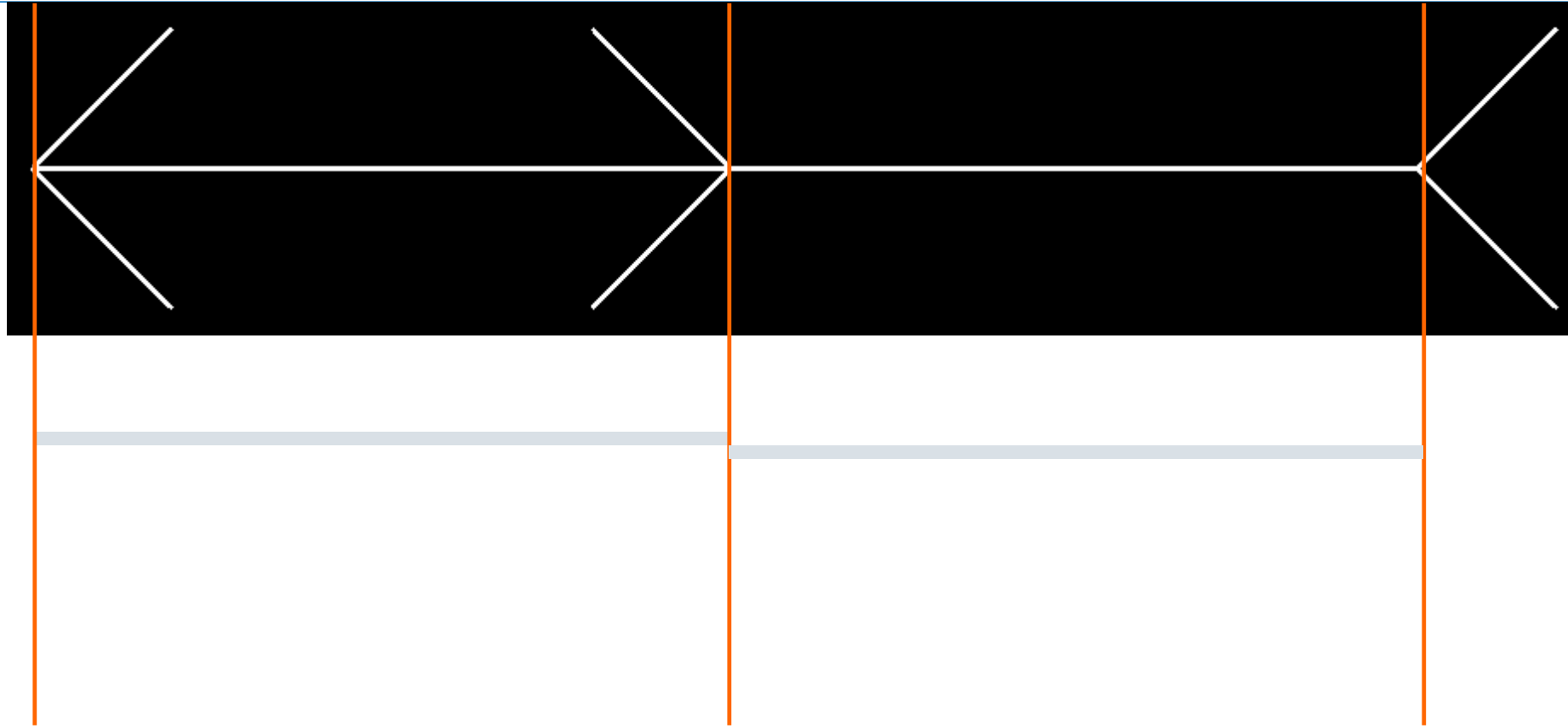


Figura Sandersa

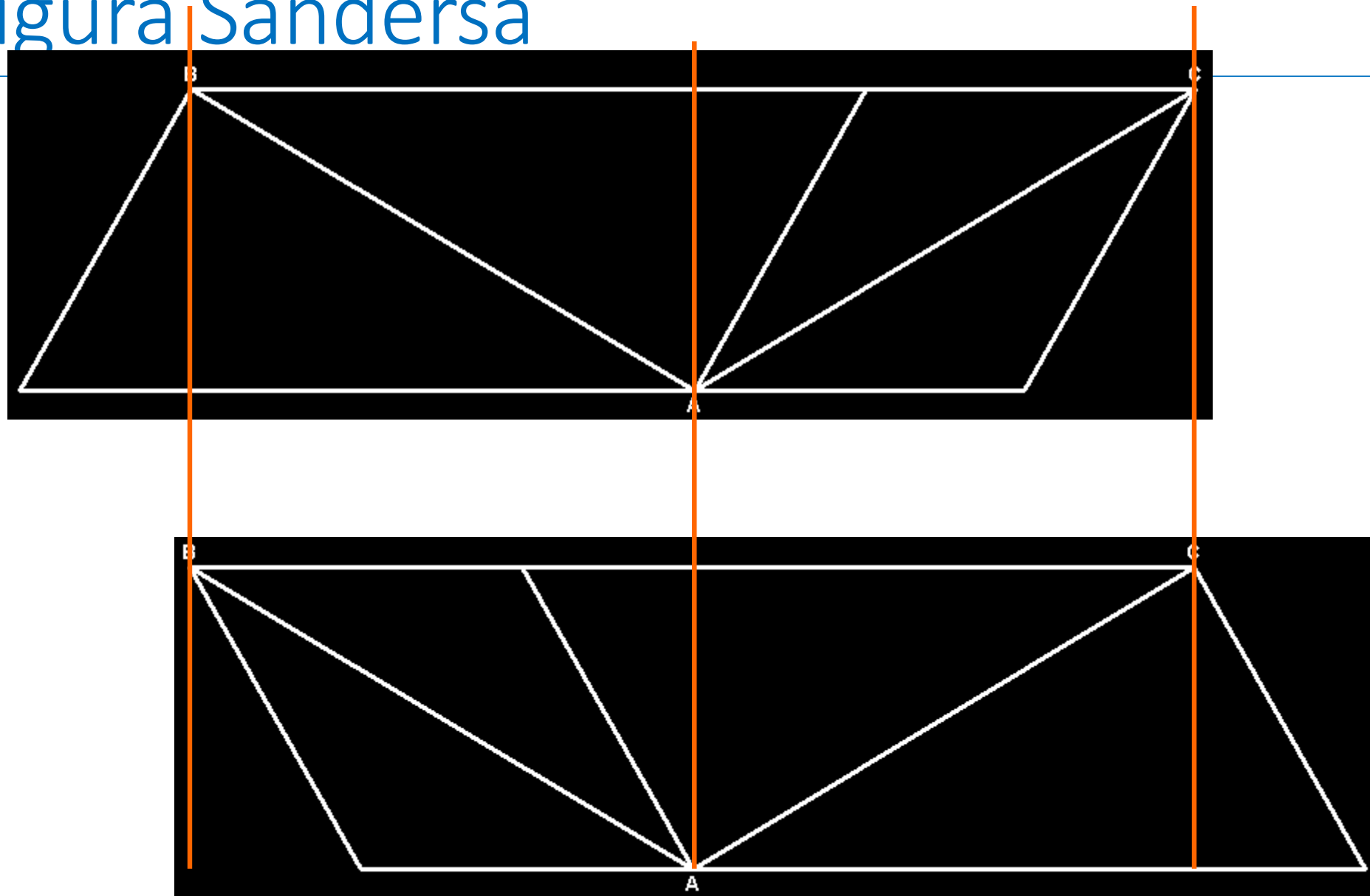


Figura Orbisona

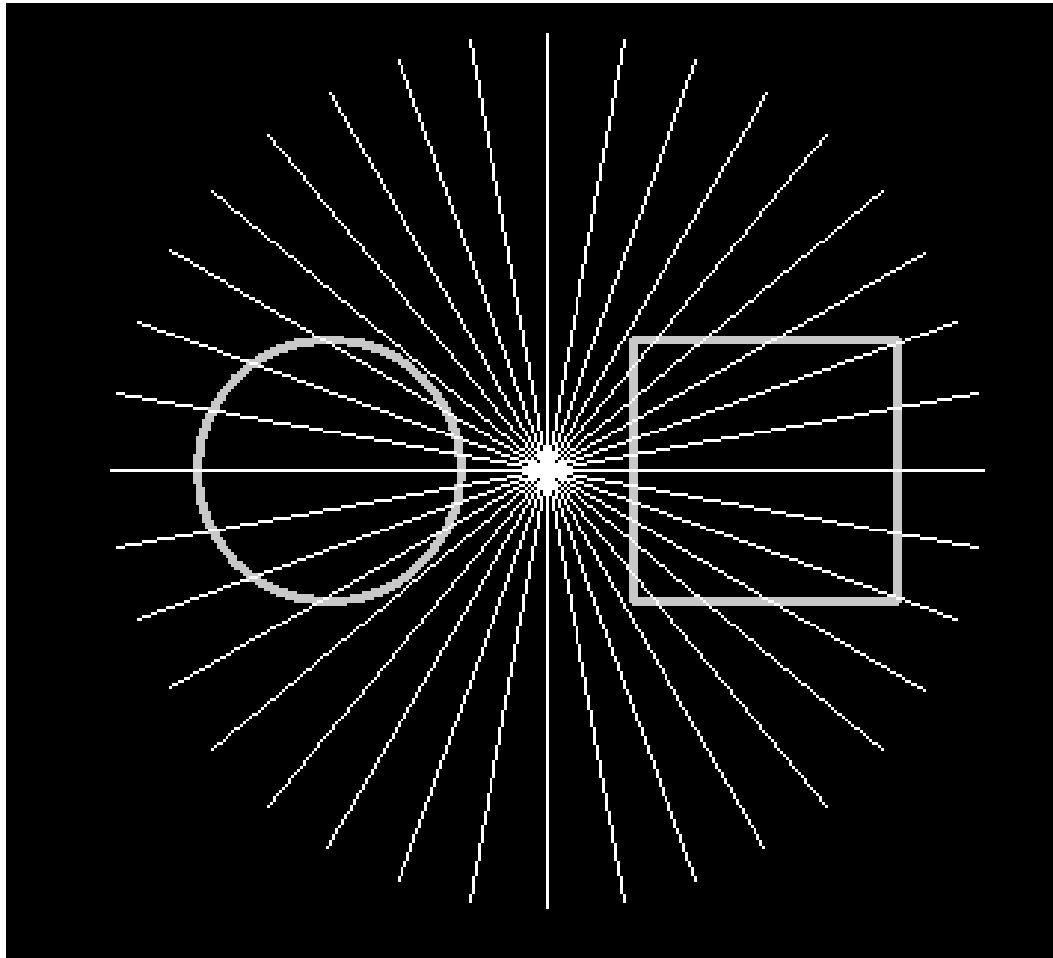
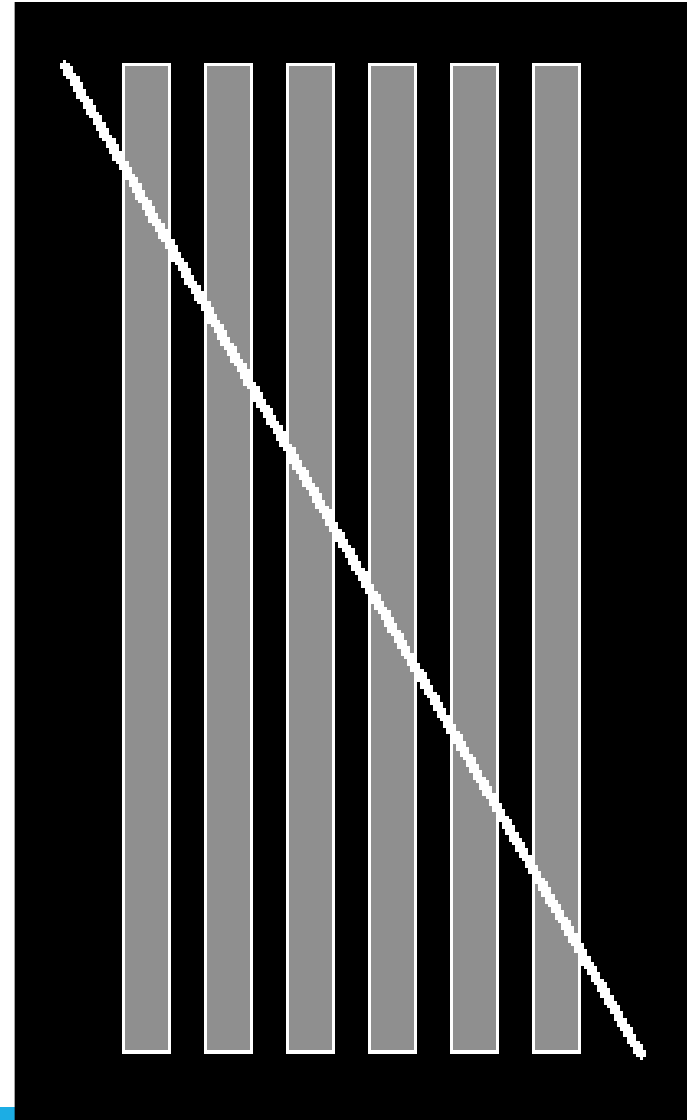
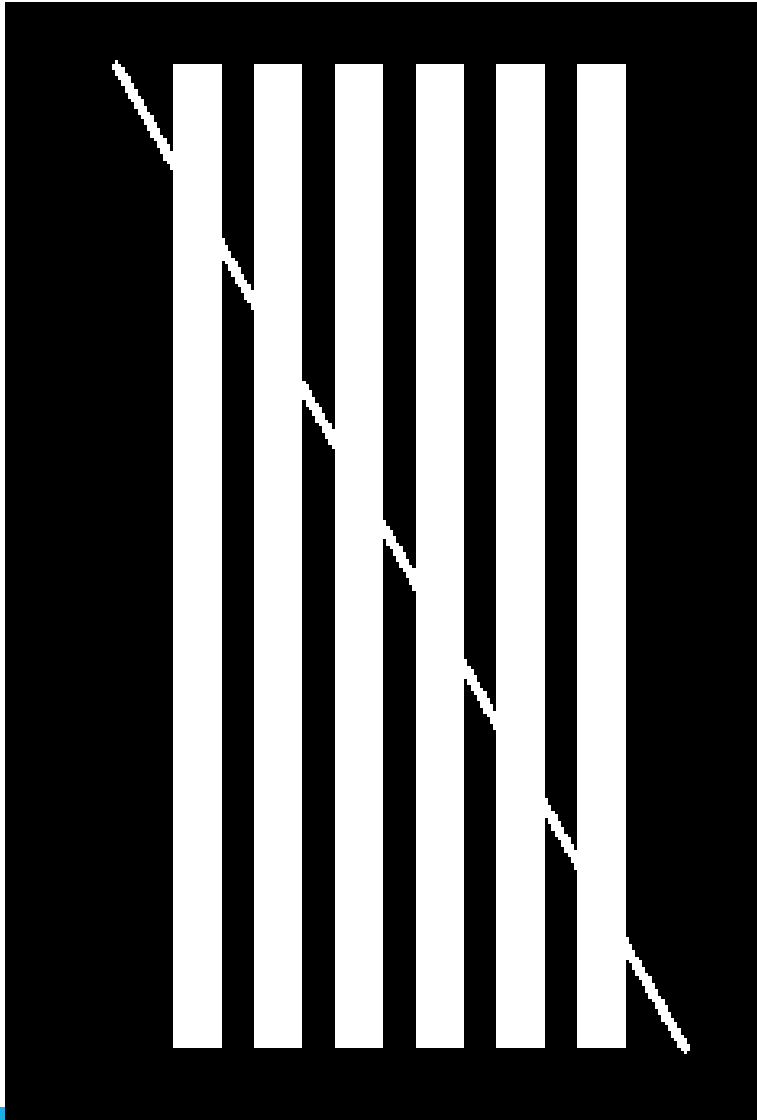
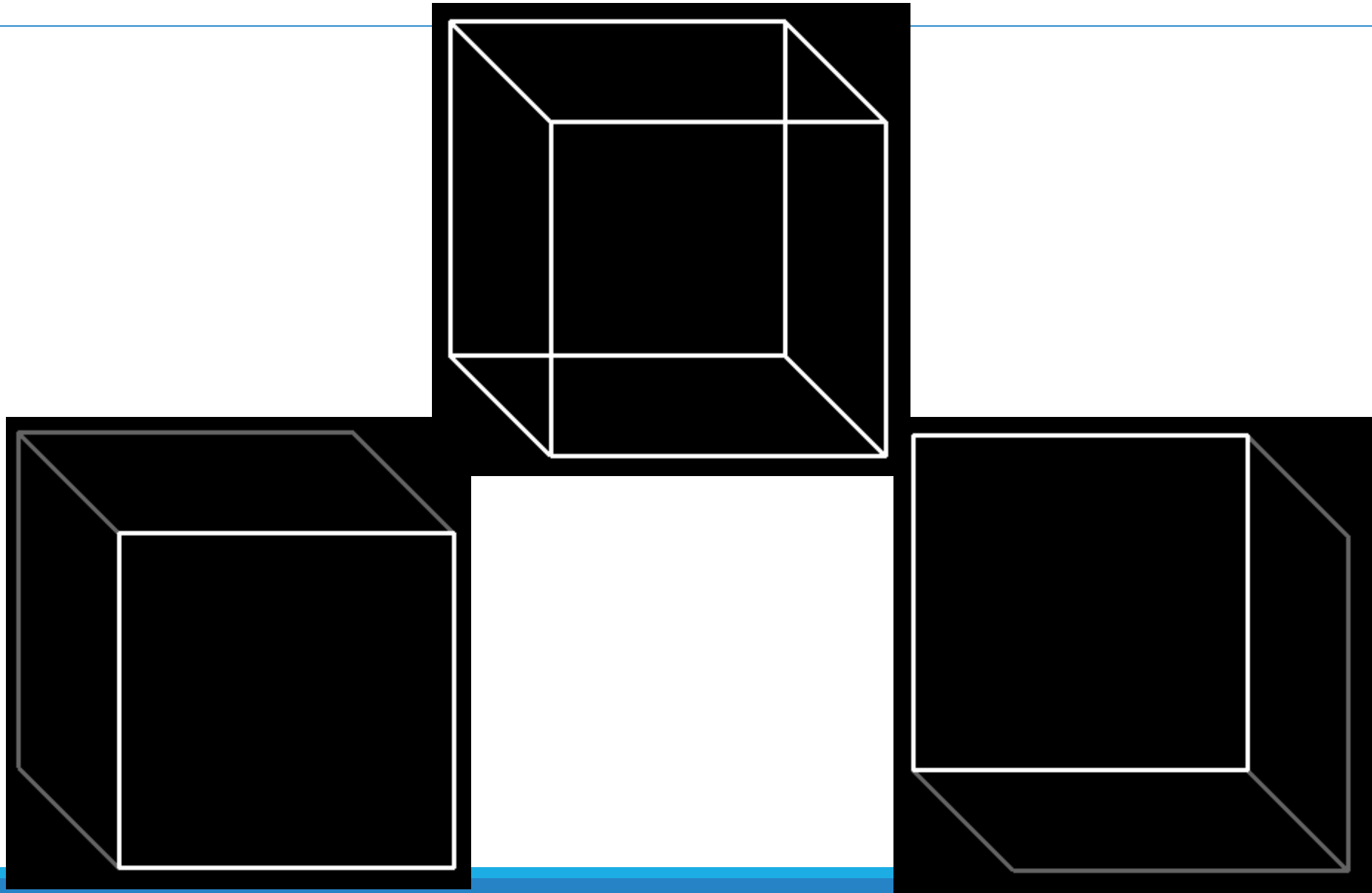


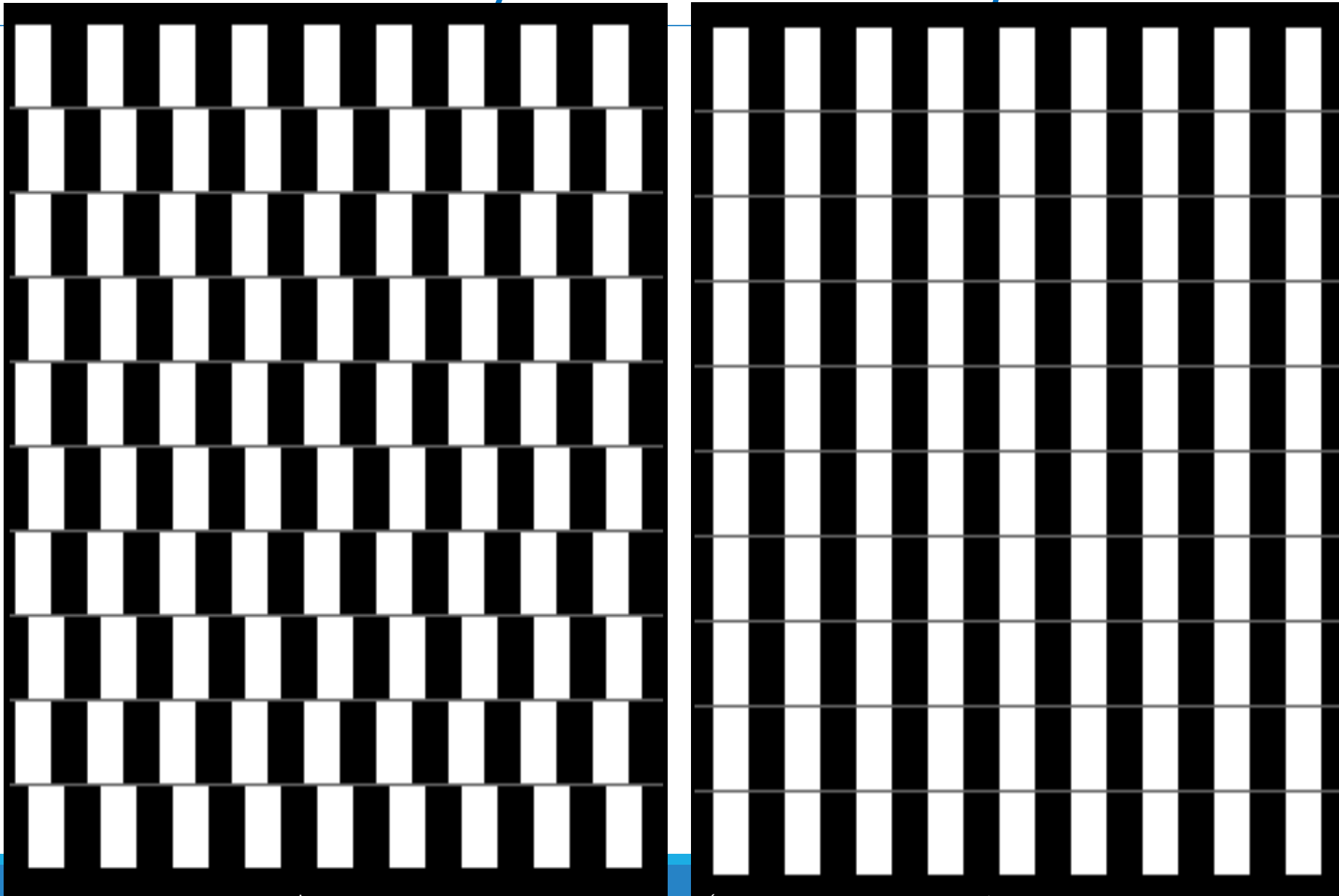
Figura Poggendorfa



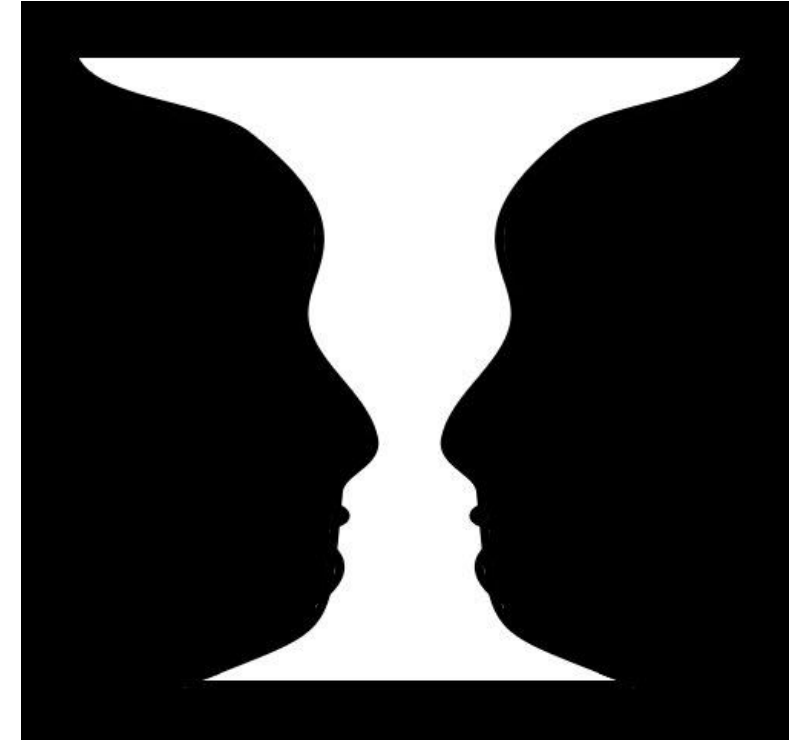
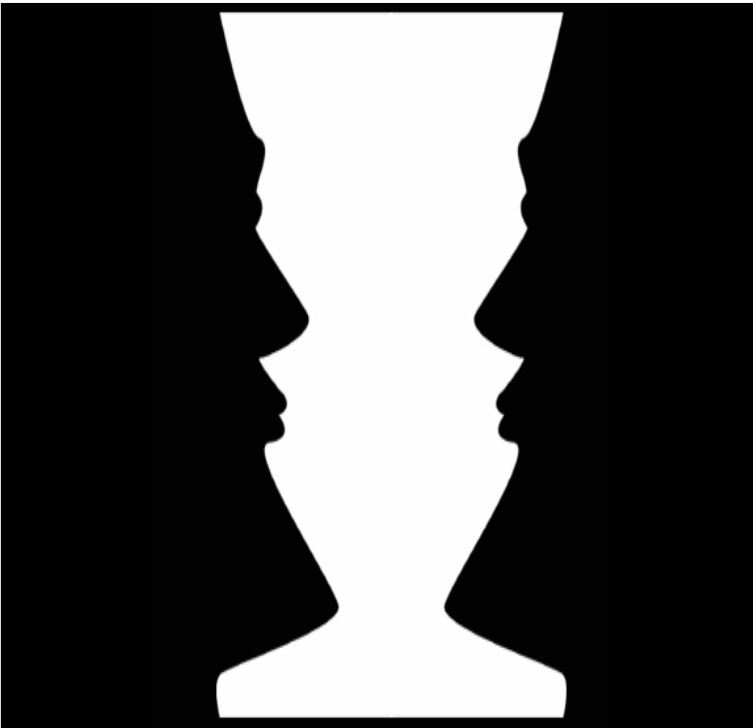
Sześćcian Neckera



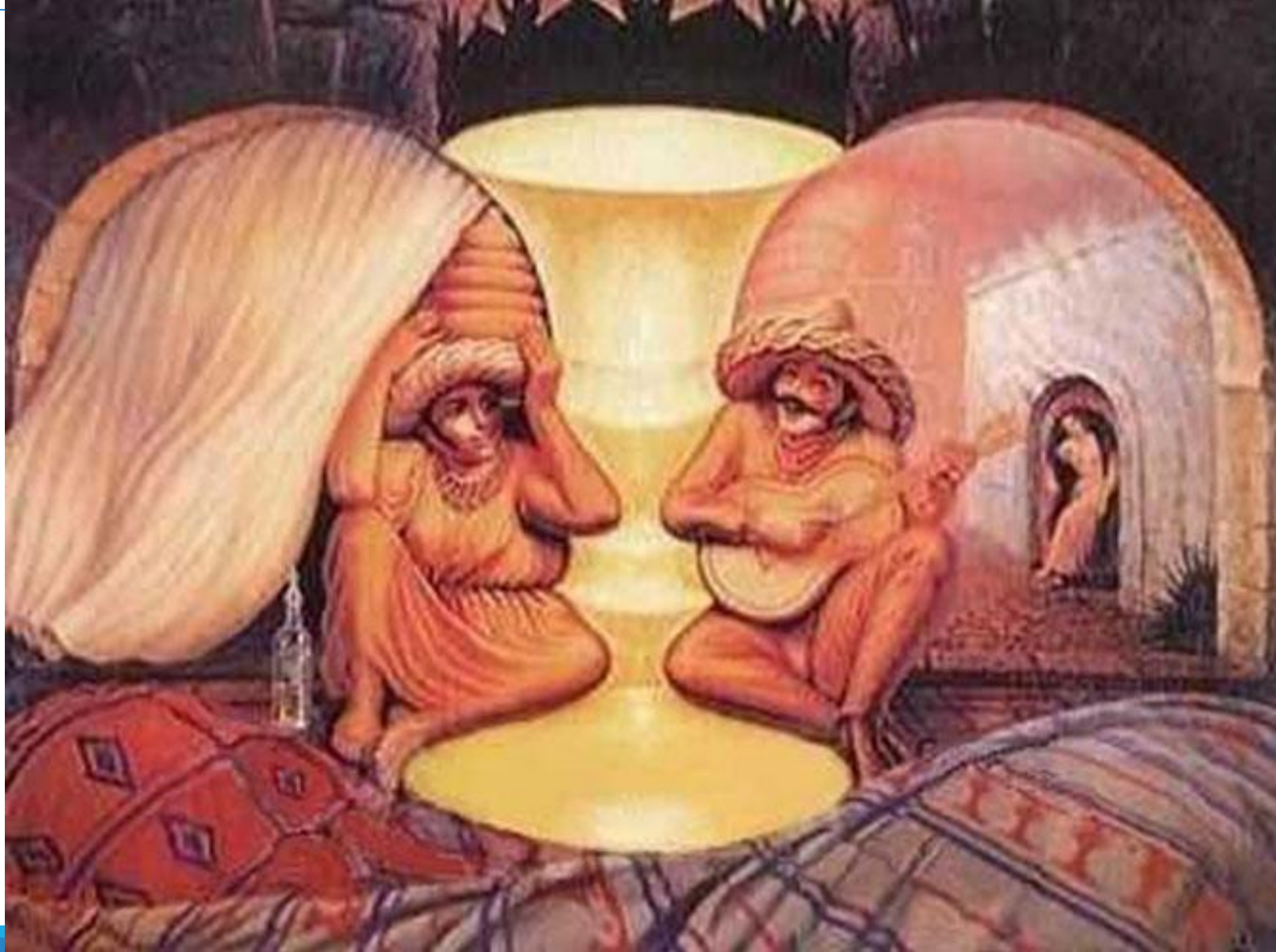
Złudzenie ściany kawiarnianej



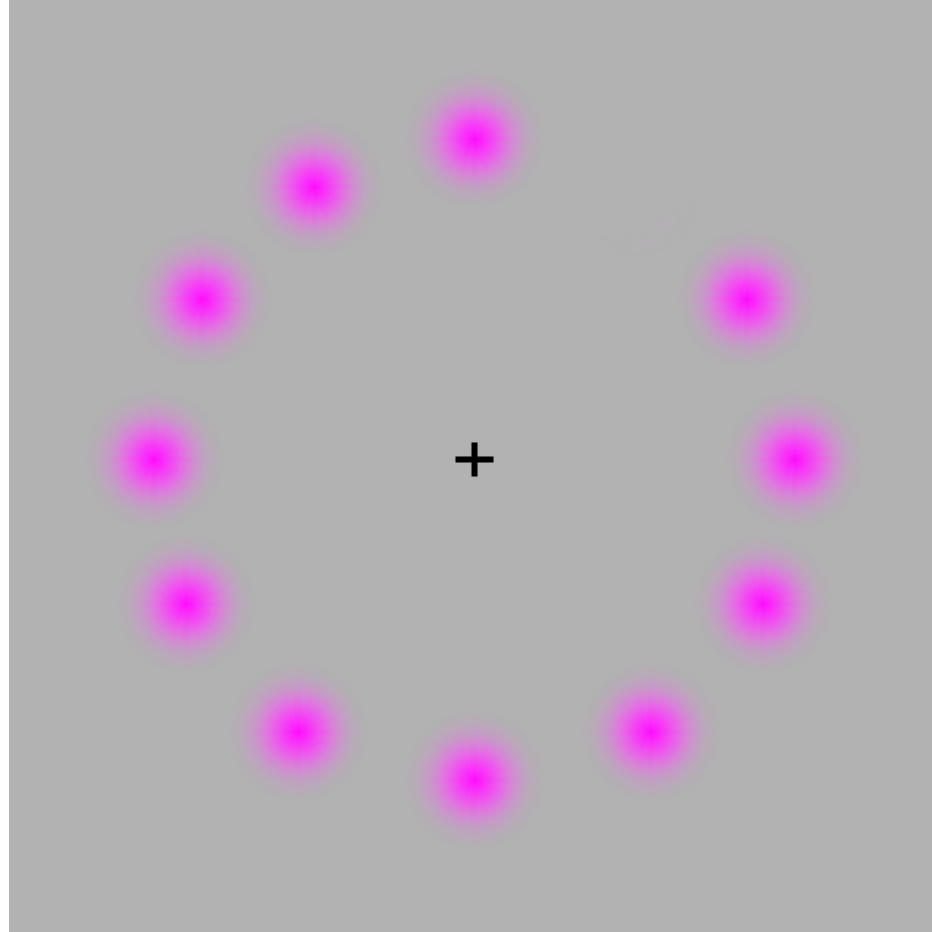
Obrazy wieloznaczne, wazon Rubina



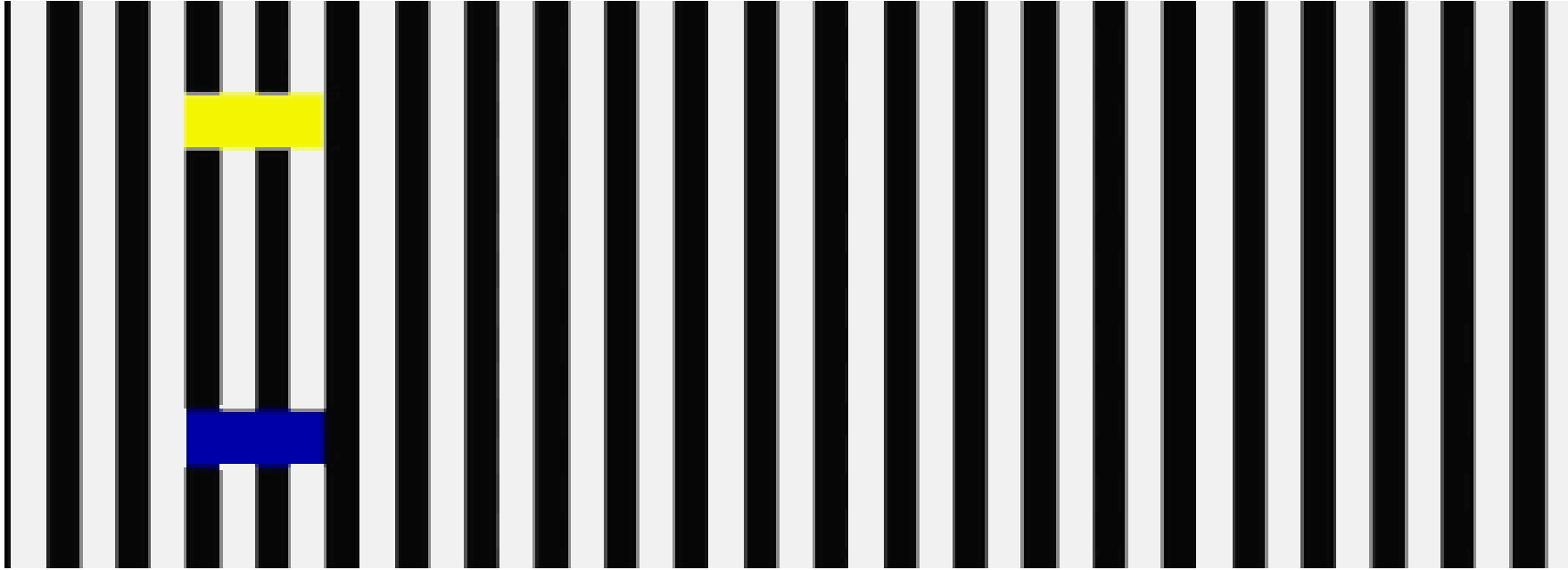
Salvador Dali



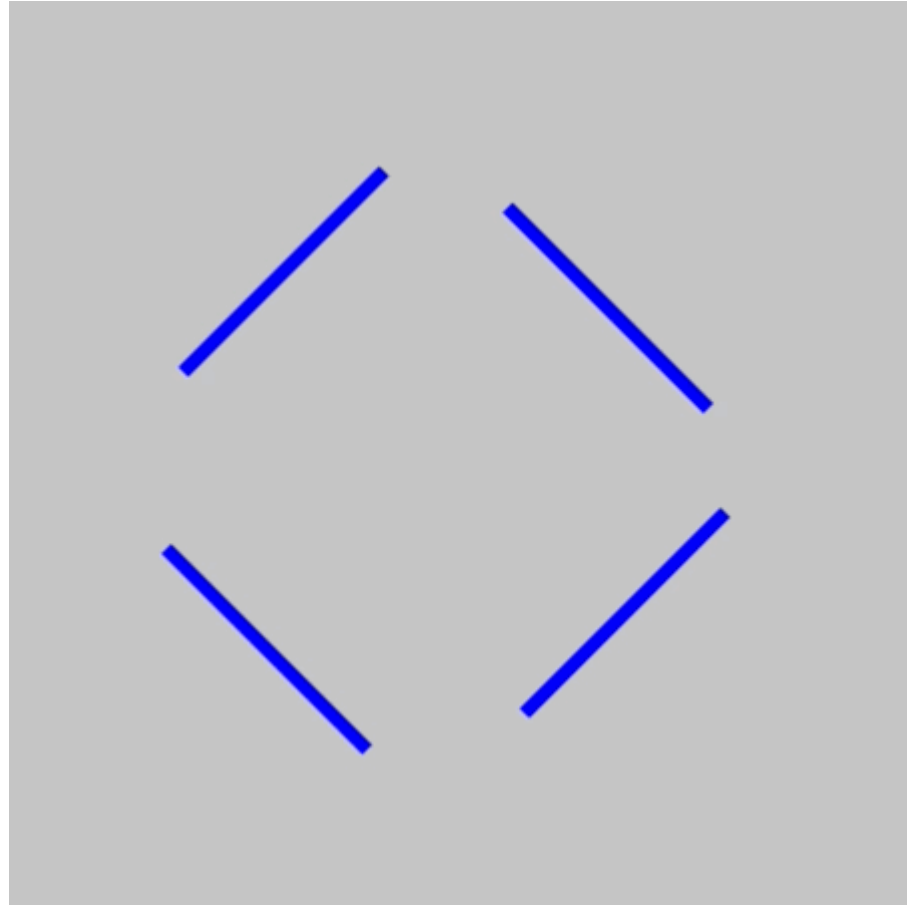
Znikające kąta



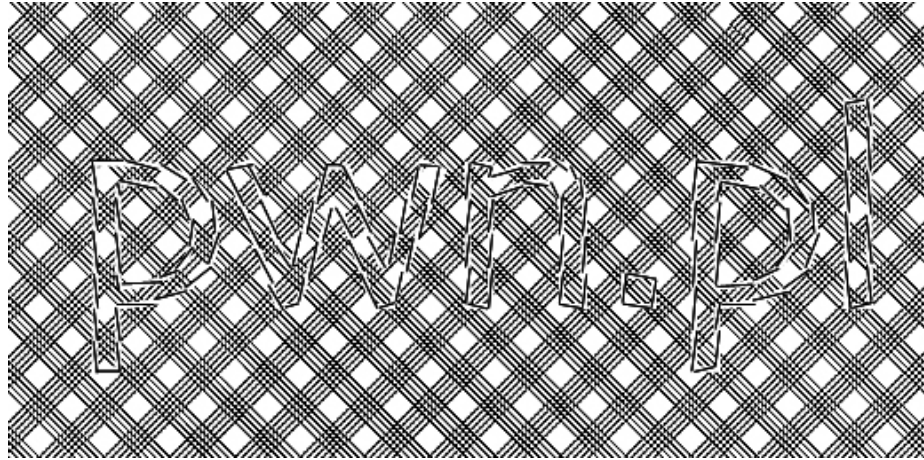
Płynność ruchu



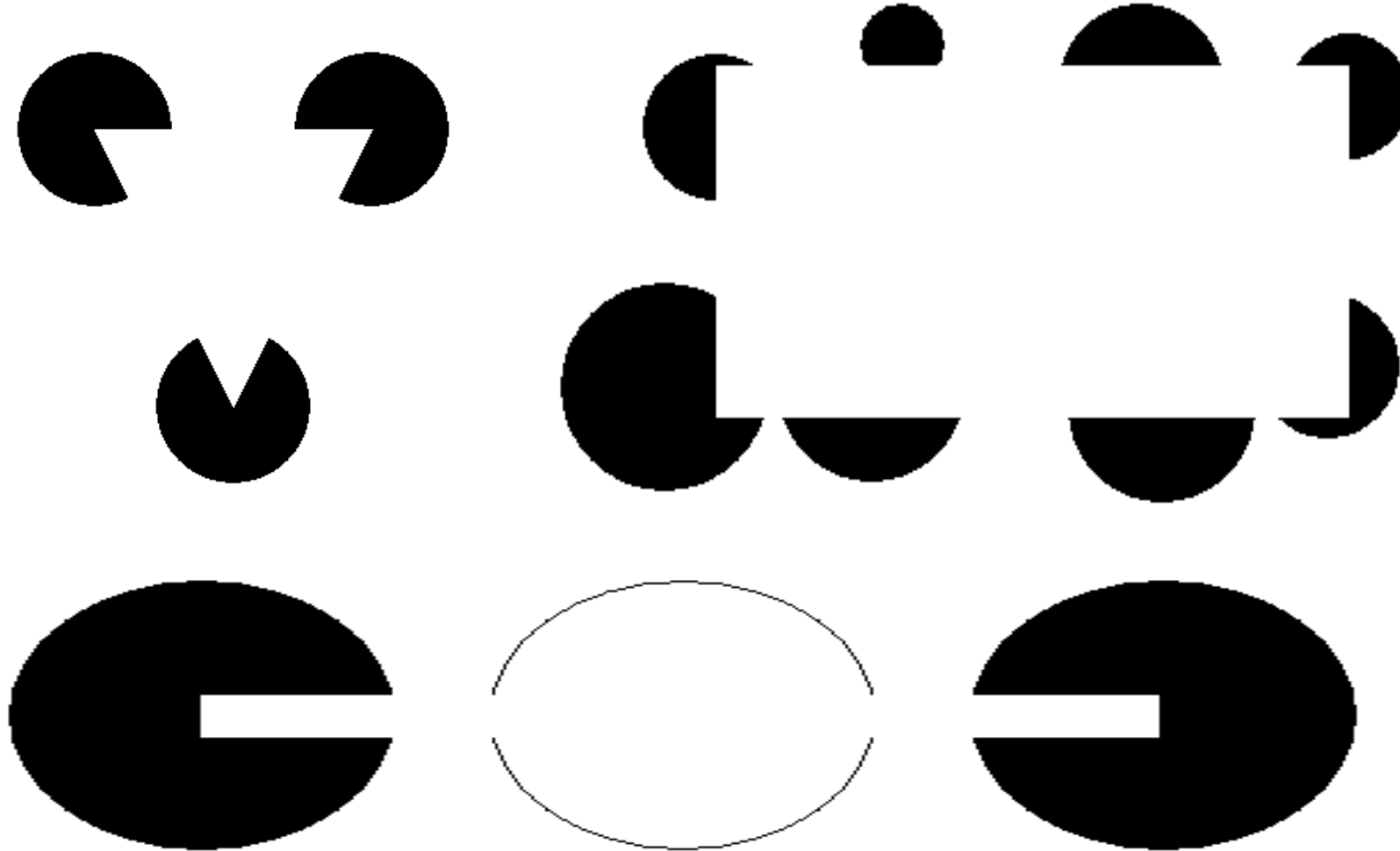
Ruchomy kwadrat



Złudzenie pochyłości

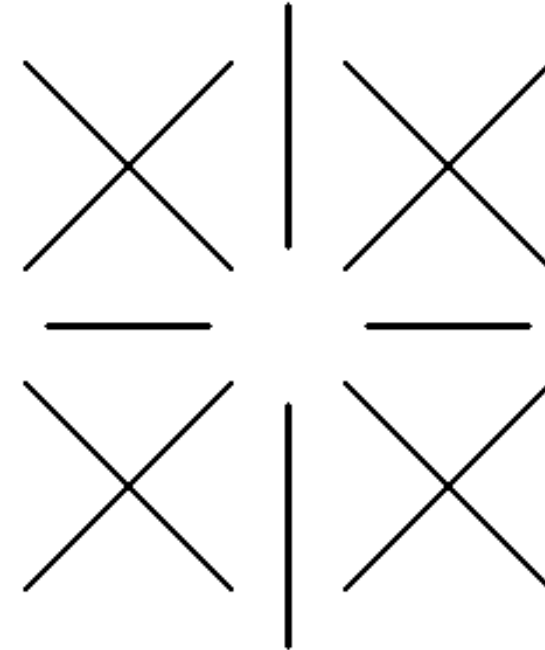
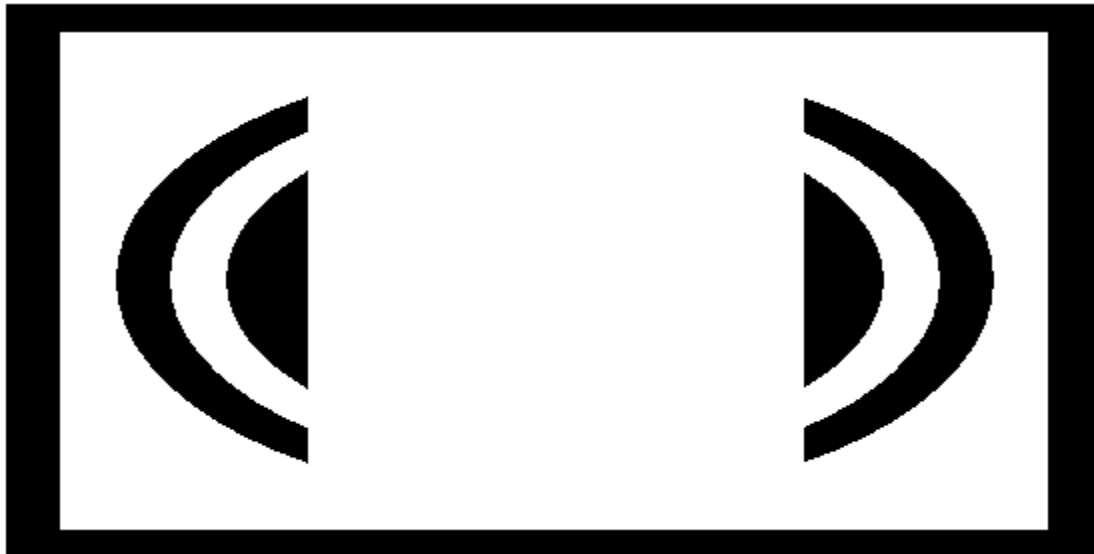


Figury iluzoryczne (1)



Figury iluzoryczne (2)

- Kwadrat Schumanna



Koło Ehrensteina

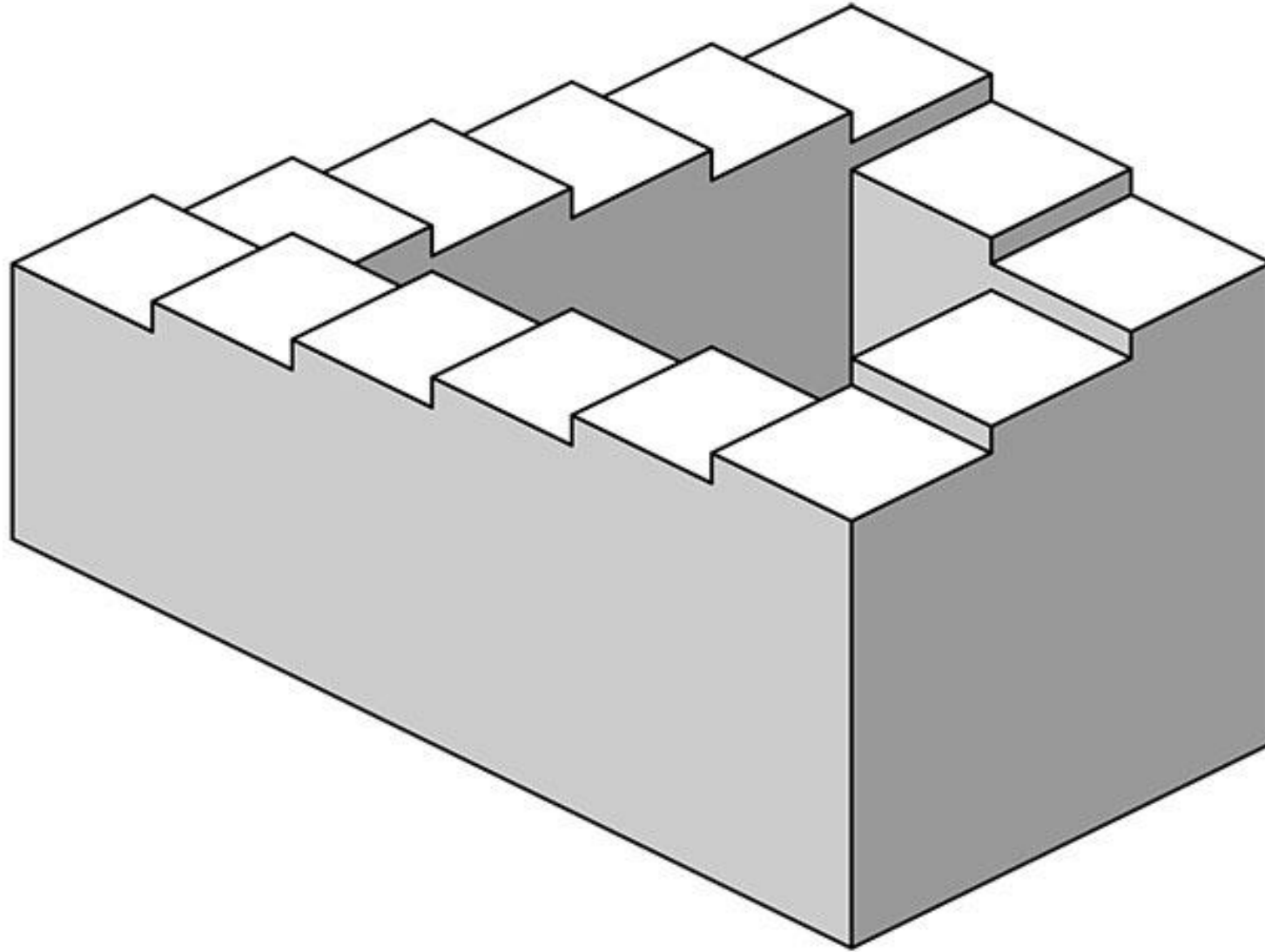
Złudzenia perspektywy (3D)



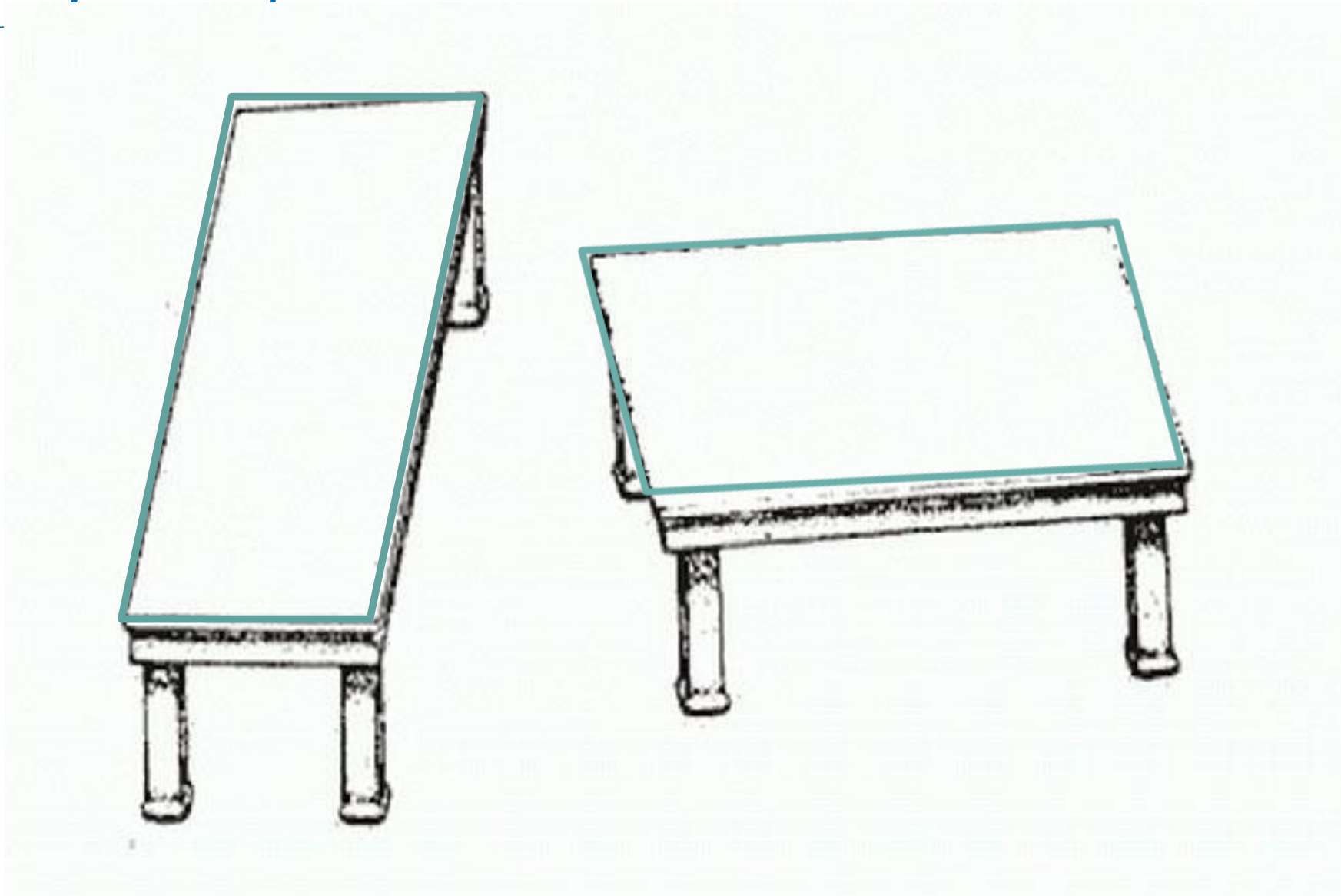
Perspektywa



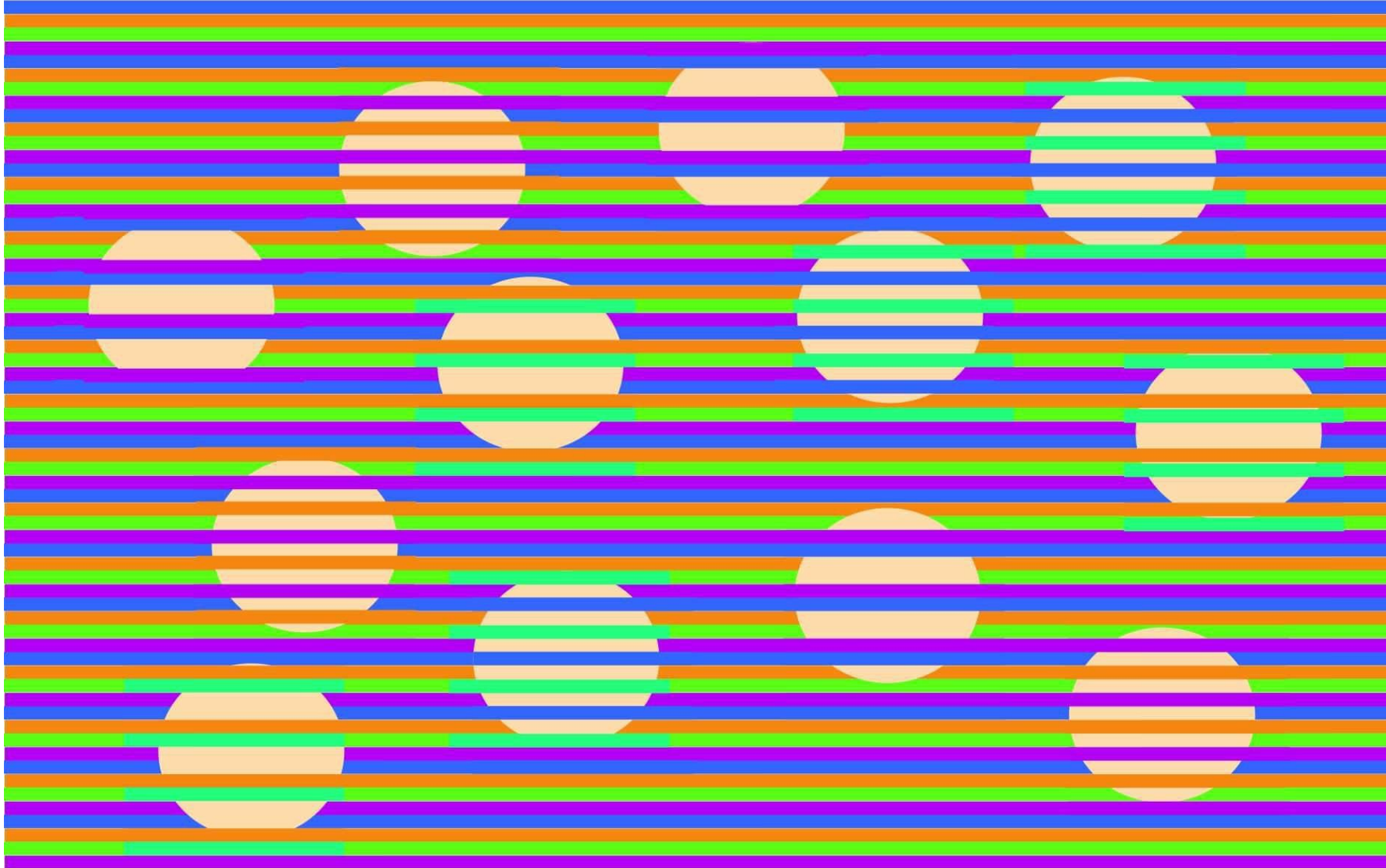
Schody Penrose'a



Stoły Sheparda



Wszystkie kropki mają ten sam kolor



Daltonizm

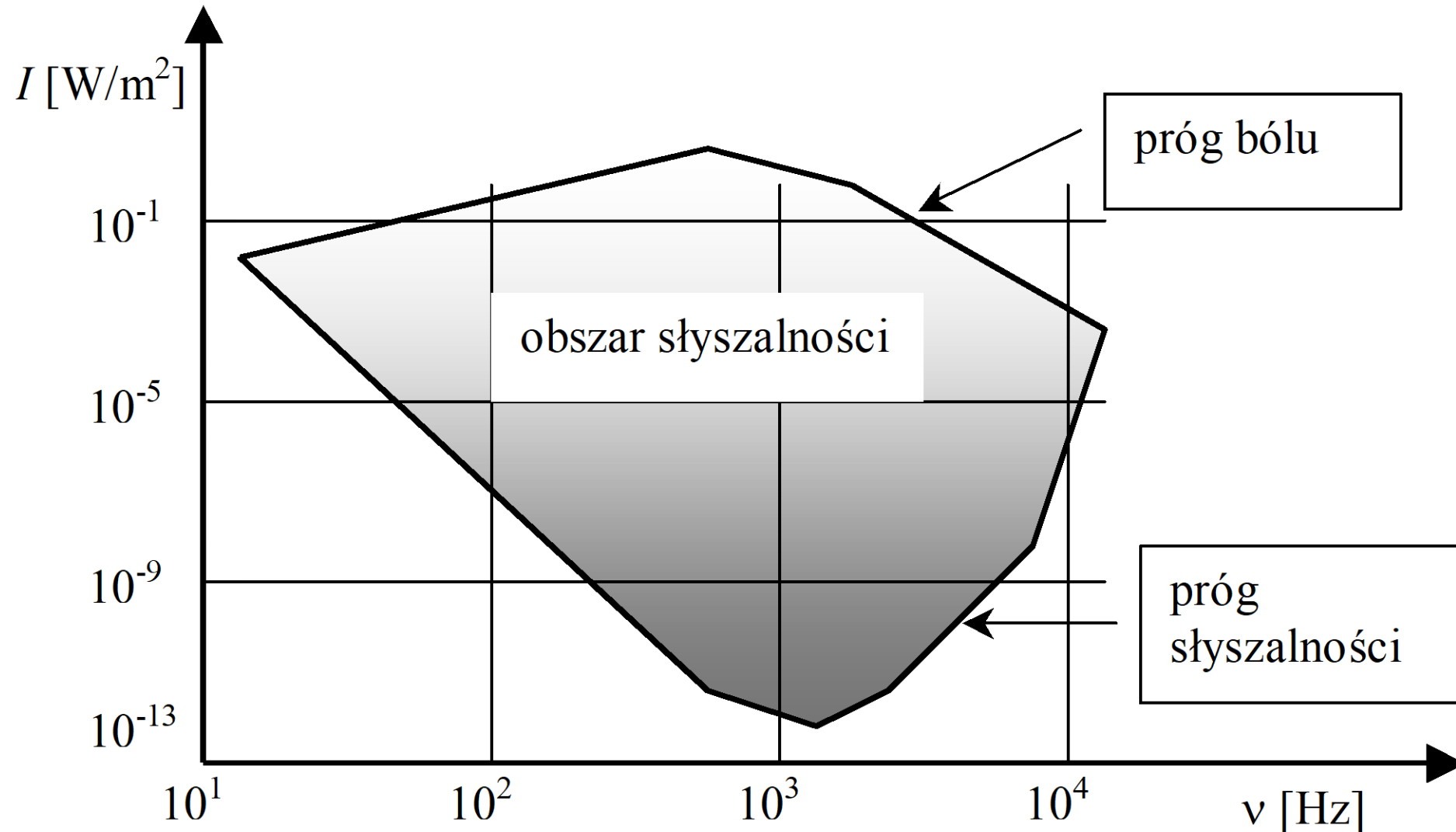
- **Daltonizm** - częściowe lub całkowite zaburzenie odbioru barw
- Wprowadź cyfrę widoczną na obrazku ;-)



1.2.2. Zmysł słuchu

- ***Wrażenia słuchowe*** - odbiór fal akustycznych rozchodzących się w ośrodku sprężystym, docierających do człowieka poprzez ucho zewnętrzne oraz drogą kostną
- ***Granice słyszalności*** dźwięku przez człowieka
 - dolna ok. 20 Hz (poniżej *infradźwięki*)
 - górna ok. 20 kHz (powyżej *ultradźwięki*)
- ***Próg słyszalności***
 - dolny $I_0 = 10^{-12} \text{ W/m}^2$ przy $\nu=3\text{kHz}$
 - górny ok. 1 W/m^2 przy $\nu=1\text{kHz}$

Krzywa czułości ucha ludzkiego

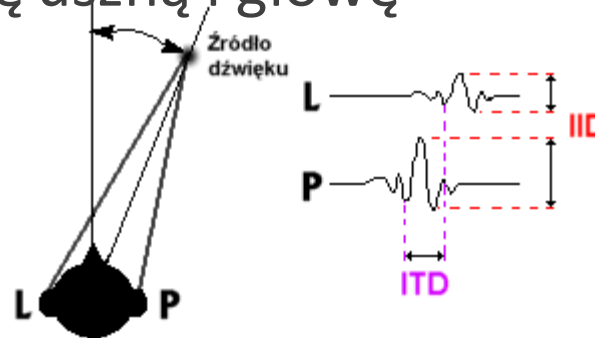


Słyszenie przestrzenne

- Różnica w czasie docierania dźwięku do uszu wynosi ok. 1 ms
- Dźwięk słyszany z tyłu jest nieco stłumiony
- Dźwięki z daleka są cichsze
- Łatwiej lokalizowane są dźwięki wyższe
- Zjawisko Dopplera zmienia częstotliwości dźwięków wydawanych przez obiekty ruchome oraz odbierane przez ruchomego słuchacza
- Człowiek rozróżnia echo i odbicia od dźwięku właściwego

Słyszenie przestrzenne

- IID – *Interaural Intensity Difference*; różnica intensywności dźwięku odbierana między uszami
- ITD – *Interaural Time Difference*; różnica czasu, w jakim dociera dźwięk
- HRTF – *Head-Related Transfer Function*; efekt tłumienia/wzmacniania dźwięku przez małżowinę uszną i głowę



Skala decybelowa

- **Ciśnienie akustyczne L** - stosunek natężenia I dźwięku do progu słyszalności I_0 przy częstotliwości 1 kHz w skali logarytmicznej

$$L = 10 \lg \frac{I}{I_0}$$

- **Decybel** - względna logarytmiczna jednostka natężenia dźwięku oraz ciśnienia akustycznego

Głośność

- **Głośność** - wielkość charakteryzująca subiektywne odczuwanie dźwięku przez człowieka; zależy od natężenia i częstotliwości dźwięku
- **Fon** - jednostka poziomu głośności wyrażona w skali decybelowej; 1 fon odpowiada poziomowi głośności dźwięku o częstotliwości równej 1 kHz i o poziomie ciśnienia akustycznego 1 dB
(1 fon = 1 dB, przy 1kHz)

1.2.3. Inne zmysły i odczucia człowieka

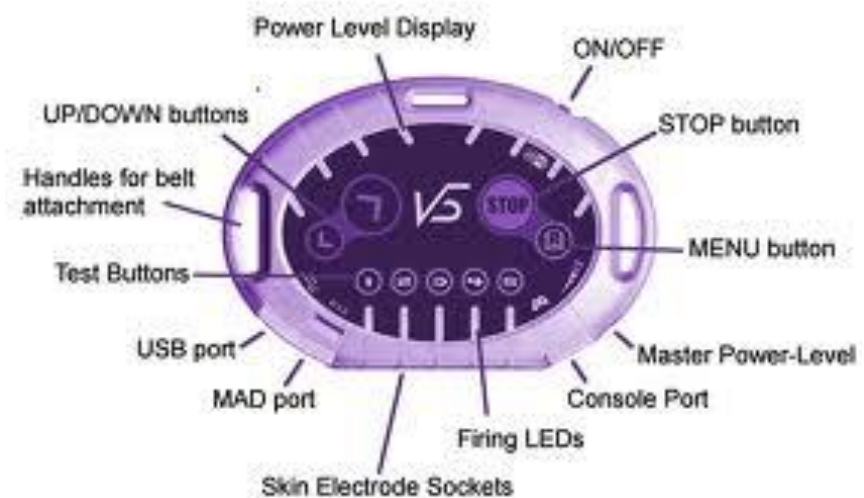
- Zmysły wrażliwe na związki chemiczne
 - węch - nos - nabłonek węchowy
 - smak - język, podniebienie - kubki smakowe
- Zmysły somatyczne (czucie skórne)
 - dotyk, nacisk - nerwy czuciowe
 - ból (nocycepja)
 - temperatura (ciepło/zimno)
- Równowaga - błędnik
- Odczucia
 - napięcie elektrostatyczne, prąd elektryczny
 - wilgotność, lepkość
 - ciśnienie

Węch, smak, dotyk, ..

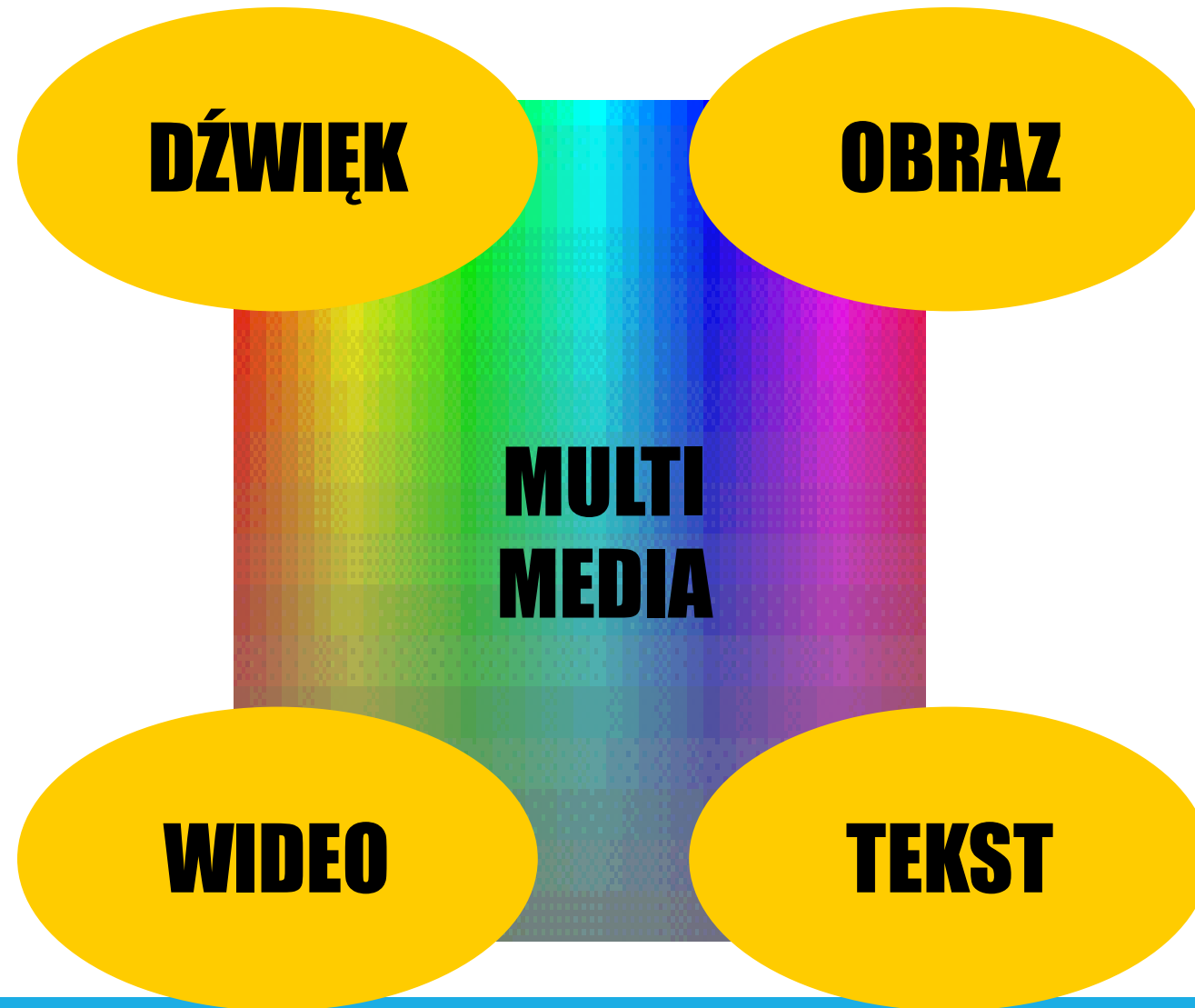
- **Smak** - 5 rodzajów receptorów związków chemicznych
 - gorzki, słodki, kwaśny, słony, umami (kwas glutaminowy)
- **Węch** - ok. 50 milionów receptorów dla ponad 100 związków chemicznych
- **Dotyk** - przenoszenie mechanicznych (tylko) wrażeń dotykowych
 - urządzenia wirtualnej rzeczywistości VR (*virtual reality*):
 - interfejsy PC ze sprzężeniem zwrotnym (wibracje)
 - kierownice, pady, joysticki
 - inne odczucia
 - konsystencja, struktura, lepkość, prąd elektr., ..

Mindwire

- Impulsy elektryczne przy sprzężeniu zwrotnym
 - PC, PS, XBox



1.3. Akwizycja danych multimedialnych



Rodzaje mediów

- Tekst
 - drukowany
 - na ekranie
 - książki elektroniczne
 - hipertekst
- Grafika
 - rastrowa
 - wektorowa
 - trójwymiarowa
- Dźwięk
 - odtwarzany
 - generowany
- Animacja
 - animacja komputerowa
 - animacja 3D
- Wideo

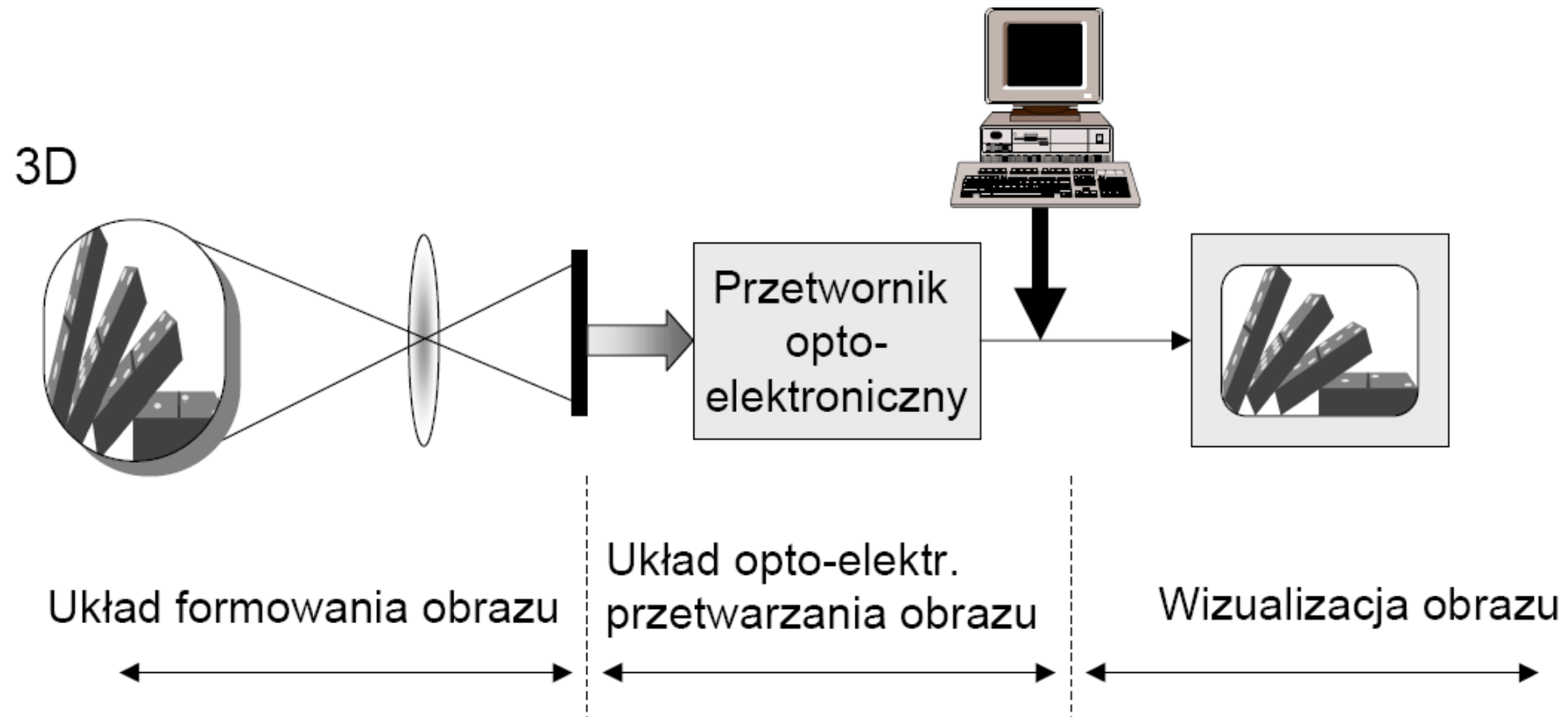
Dyskretyzacja i kwantyzacja w 4D

- **Sygnal analogowy** – zjawisko fizyczne ciągłe w przestrzeni i w czasie charakteryzowane przez pewne wielkości analogowe
- **Pomiar sygnału** – określenie wartości parametrów sygnału
- **Digitalizacja** – konwersja sygnału analogowego na cyfrowy
- **Dyskretyzacja (próbkiowanie)** - określenie punktów pomiaru sygnału (w czasoprzestrzeni)
- **Kwantyzacja** - przedstawienie wartości parametru sygnału z określoną (skończoną) dokładnością

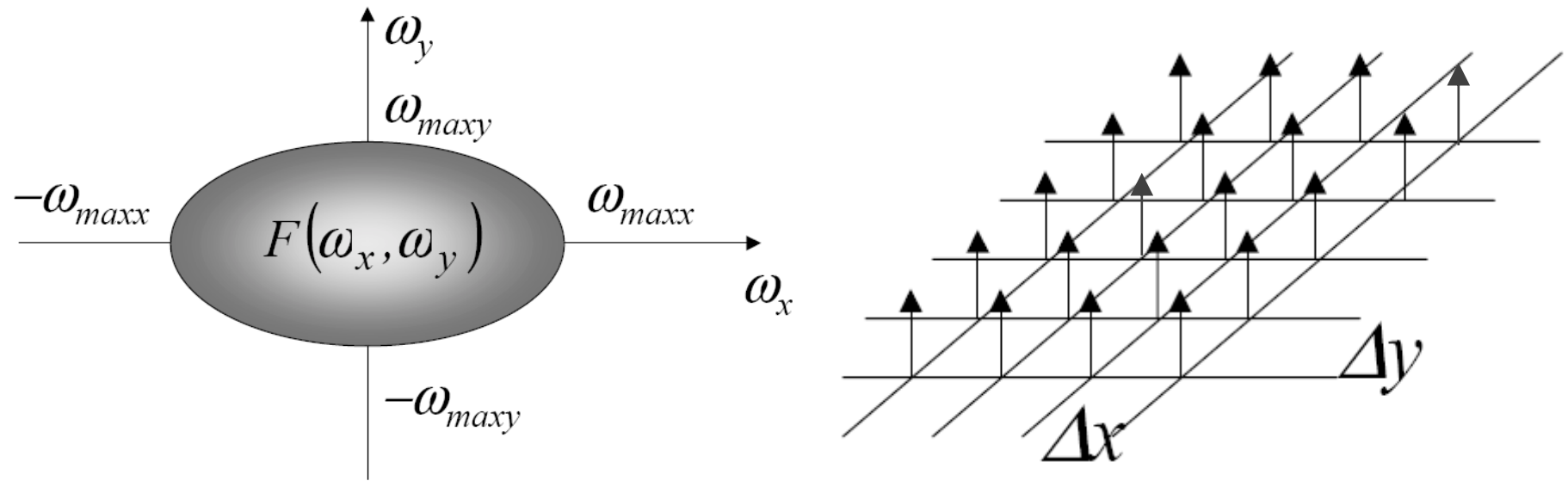
1.3.1. Akwizycja obrazów statycznych

- **Dyskretyzacja** - wyznaczenie wartości koloru obrazu ciągłego w wybranych miejscach w danym momencie czasu:
 - siatka prostokątna, heksagonalna, inne
- **Rozdzielczość dyskretyzacji** - liczba punktów pomiarowych na 1 cal długości (dpi - *dots per inch*)
- **Kwantyzacja** - zaokrąglenie poszczególnych wartości koloru w każdym punkcie pomiaru do najbliższej wartości dyskretnej
- **Rozdzielczość kwantyzacji** (głębia kolorów) - liczba bitów przeznaczonych do zapisu każdego punktu dyskretnego (bpp - *bits per pixel*)

Tor wizyjny systemu przetwarzania obrazów



Dyskretyzacja obrazu



$$S(x, y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(x - i\Delta x, y - j\Delta y)$$

$$f_s(x, y) = f(x, y)S(x, y) = S(x, y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i\Delta x, j\Delta y) \delta(x - i\Delta x, y - j\Delta y)$$

Dyskretny obraz 2D - mapa bitowa

- **Mapa bitowa I** - tablica o wymiarach $N \times M$, której elementy zawierają informację o kolorze obrazu
- **Piksel $P(i,j)$** - pojedynczy element bitmapy I
- **Paleta kolorów** - tablica kolorów występujących (dozwolonych) na bitmapie (piksel zawiera wówczas indeks do palety)
- **Rozdzielczość** - liczba punktów na cal długości obrazu w pionie i poziomie (DPI - *dots per inch*)

Głębia kolorów

- Rozmiar opis koloru
 - **1 bit** - obraz czarno-biały B/W (*black and white*)
 - 4 bity - 16 kolorów
 - **8 bitów** - achromatyczny (*achromatic*), poziomy szarości (*gray levels*)
 - 8 bitów - 256 kolorów (paleta)
 - 16 bitów - 65536 kolorów (*HighColor*)
 - **24 bity** - 16 mln kolorów *True Color* (RGB)
 - **32 bity** - *True Color* + kanał przezroczystości α (RGBA)

Rozmiar bitmapy

$$S = L_x R_x L_y R_y \left\lceil \frac{\log_2 C}{8} \right\rceil \quad [B]$$

S - rozmiar bitmapy w bajtach

L_x, L_y - wymiary obrazu [cm]

$$S = M \cdot N \cdot D \quad [B]$$

R_x, R_y - rozdzielczość bitmapy [DPI]

C - liczba kolorów , $\log_2 C$ - liczba bitów

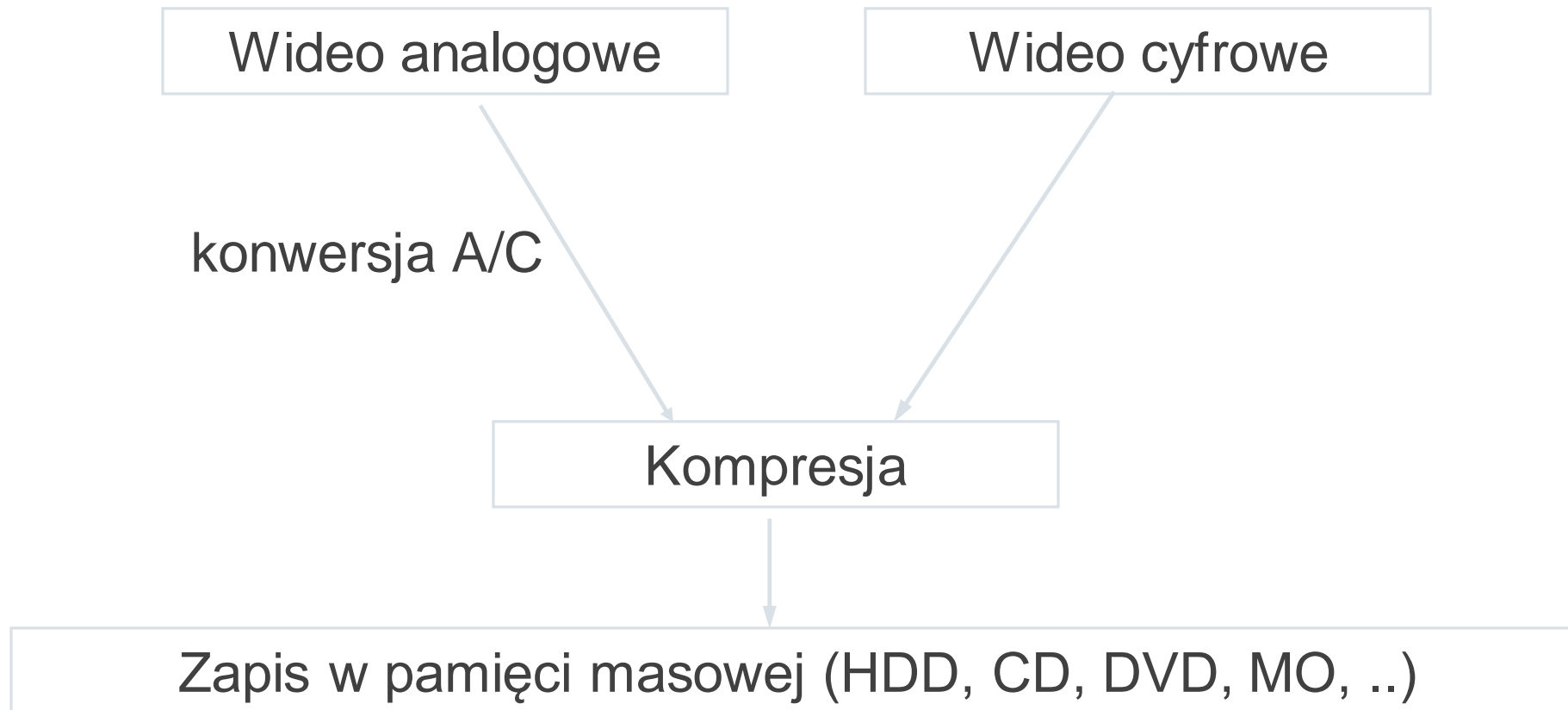
Przykład: zdjęcie paszportowe na piksel

wymiary: 3.4 cm x 4.4 cm = 1 1/3 " x 1 3/4 "

	100 dpi	300 dpi	600 dpi	1200 dpi
256 kolorów	22,7 kB	204,5 kB	820,3 kB	3,14 MB
64k kolorów	45,4 kB	409 kB	1,64 MB	6,4 MB
16mln kolorów	68,1 kB	613,5 kB	2,4 MB	9,6 MB

1.3.2. Akwizycja obrazów dynamicznych

- **Wideo** - sekwencja obrazów statycznych
- Dyskretyzacja w dziedzinie czasu (próbkiowanie) i przestrzeni



Wideo - tablica ramek

- **Tablica ramek** - jednowymiarowa tablica której elementami są poszczególne ramki (klatki) obrazu wideo zapamiętane w kolejnych chwilach czasu
- **Ramka obrazu** (klatka) - mapa bitowa zawierająca nagrany obraz
- **Częstotliwość nagrania** - liczba ramek zapisanych w ciągu sekundy (fps - *frames per second*)

Typowe parametry wideo

- Typowe rozmiary ramki wideo:
 - telewizja, wideo analogowe, PC
 - 352×288 ramka TV, VHS, kamera internetowa
 - 320×576 PAL VHS
 - 640×480 kamera internetowa, VGA
 - 720×576 PAL
 - telewizja cyfrowa
 - 720×486 PAL (720×576 D-1 PAL, 768×576 D-1 PAL)
 - 1920×1080 (1280×720) HDTV, 4k, 8k
 - film cyfrowy
 - 720×576 DVD (720×480)
 - 2048×1536 Standard akademii filmowej, 4k, 8k
- Typowe częstotliwości nagrań wideo:
 - 15-25 fps - kamera internetowa, 25 fps - PAL, 30 fps - NTSC

Zajętość pamięci

$$S = N \cdot M \cdot F \cdot T \cdot \left\lceil \frac{\log_2 C}{8} \right\rceil [B]$$

S - rozmiar tablicy ramek w bajtach

N, M - rozmiar ramki [piksel]

F - częstotliwość nagrania [fps]

T - czas nagrania [s]

C - liczba kolorów , $\log_2 C$ - liczba bitów

Przybliżony rozmiar 1 min. nagrania bez kompresji: na piksel

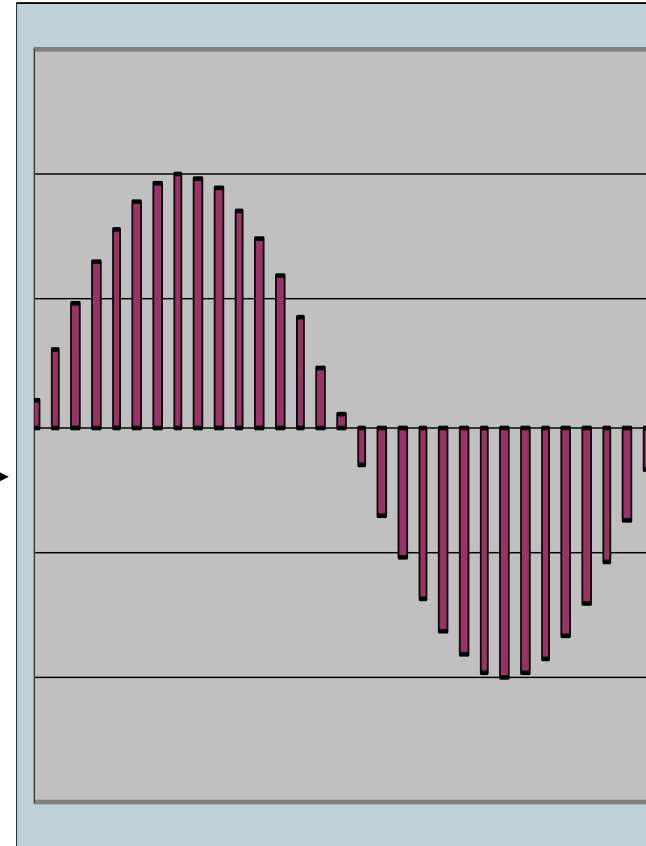
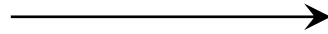
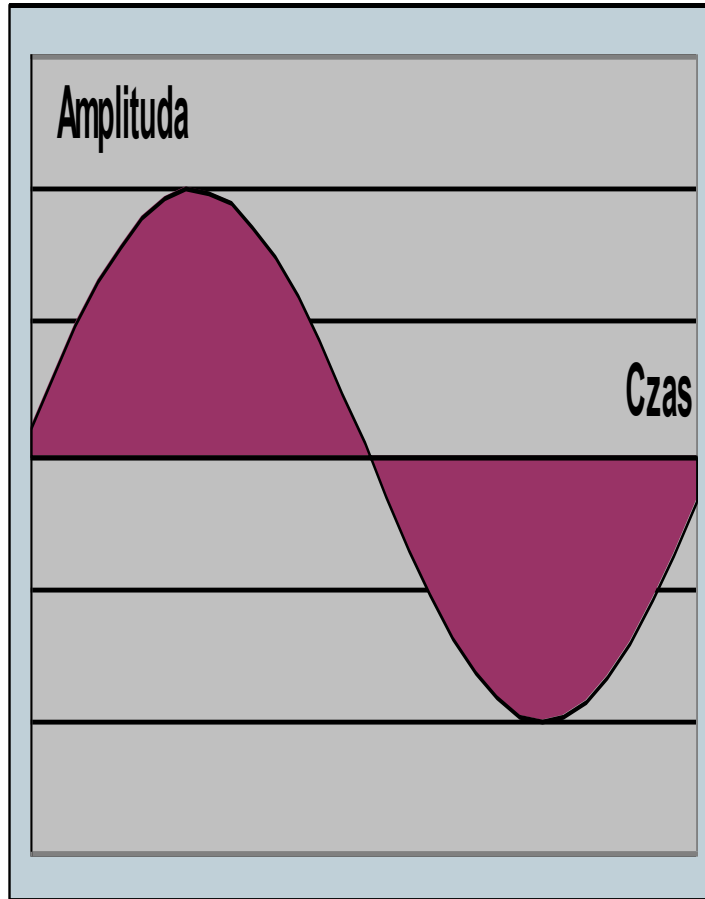
- PAL - 1,8 GB
- NTSC - 1,5 GB
- HDTV - 5 GB

1.3.3. Akwizycja dźwięku

- **Próbkowanie** - proces wyznaczania wartości sygnału (kwantyzacja) w określonych momentach czasowych (dyskretyzacja)
- **Częstotliwość próbkowania** (*sample rate*) - liczba próbek pobierana z sygnału analogowego w jednostce czasu (Hz)
- **Twierdzenie Shannona:**

Częstotliwość próbkowania musi być co najmniej dwa razy wyższa od najwyższej częstotliwości w próbkowanym sygnale, którą chcemy później odtworzyć

Próbkowanie dźwięku



Parametry próbkowania

- **Częstotliwość próbkowania** (*sample rate*) - liczba próbek na sekundę (Hz)
- **Rozdzielczość kwantyzacji** - liczba bitów na próbkę (*bits*):
 - 8, 16, 24, 32
- **Liczba kanałów** (dyskretyzacja przestrzenna)
 - 1 (mono), 2 (stereo), 4, 5+1, 6+1, 7+1

Dźwięk cyfrowy – dynamika

- **Dynamika sygnału** – odstęp pomiędzy najniższym i najwyższym poziomem przenoszonego dźwięku
- **SNR** (*Signal to Noise Ratio*) - stosunek mocy sygnału użytecznego do szumu [dB]
- Przybliżony wzór określający zależność dynamiki sygnału cyfrowego od rozdzielczości kwantyzacji:

$$D = R * 6 + 1.76 \text{ [dB]},$$

D - dynamika, R – rozdzielczość kwant.

Przykładowo:

- 16-bitów - ok. 98 dB,
- 8-bitowe - 50 dB

Popularne częstotliwości próbkowania

- 8 kHz – telefonia, sygnał mowy
- 18.9 kHz – Standard CD-ROM/XA.
- 32 kHz – radio cyfrowe, NICAM (*Nearly Instantaneous Compandable Audio Matrix* [IBA/BREMA/BBC]), long play DAT, japońska HDTV
- 37.8 kHz – CD-ROM/XA.
- **44.1** kHz – płyta CD
- 48 kHz – taśma DAT (Digital Audio Tape)
- 96 kHz – profesjonalny sprzęt, inżynieria dźwięku
- 192 kHz – DVDaudio

Dlaczego stosowane są wysokie częstotliwości próbkowania muzyki?

Rozmiar tablicy próbek

$$S = C \cdot F \cdot T \cdot \left\lceil \frac{R}{8} \right\rceil \quad [B]$$

S - rozmiar tablicy próbek w bajtach

C - liczba kanałów

F - częstotliwość próbkowania [Hz]

T - czas nagrania [s]

R - rozdzielczość próbkowania w bitach

Przykładowe rozmiary 1 min nagrania:

8 kHz	8bit	mono	469 kB
		stereo	938 kB
22.05 kHz	8bit	mono	1 292 kB
		stereo	2 584 kB
44.1 kHz (CD audio)	16bit	mono	5 168 kB
		stereo	10 336 kB