
Latency

Wprowadzenie

Opóźnienie obrazu (*latency*) w systemach robotycznych i wizyjnych to całkowity czas, jaki upływa od fizycznego zdarzenia w świecie rzeczywistym do momentu podjęcia akcji zwrotnej. Niezależnie czy kontrolerem jest automat, czy człowiek, możemy rozbić *latency* na czas w którym informacja jest dostarczana do kontrolera oraz czas, w którym jest przyjęta i zostaje wypracowywana decyzja o rodzaju akcji.

Spis treści

WPROWADZENIE	1
ANATOMIA OPÓŹNIENIA OBRAZU (END-TO-END LATENCY PIPELINE)	2
Systemy w pełni automatyczne (Machine Vision & Robotyka)	3
Systemy z człowiekiem-operatorem (Human-in-the-Loop)	3
Porównanie rygorów czasowych i progów tolerancji	4
LATENCY W SYSTEMACH Z OPERATOREM	5
Glass-to-Glass Latency (Opóźnienie Techniczne)	5
Opóźnienie kamery (Akwizycja i odczyt): 1,5 ms – 35 ms	5
Interfejs i Transport: 1 ms – 120 ms+	6
Komputer (Host / Przetwarzanie i System): 2 ms – 40 ms	6
Monitor i Wyświetlanie: 1 ms – 50 ms	6
Human Response Time (Opóźnienie Biologiczne)	7
Percepcja Wzrokowa (Visual Processing): ~20 ms – 40 ms	7
Przetwarzanie Kognitywne i Decyzja: ~50 ms – 150 ms+	7
Transmisja Motoryczna (Motor Efferent Pathway): ~10 ms – 20 ms	7
Fizyczna Aktywacja Mięśnia (Muscle Activation & Mechanical Lag): ~30 ms – 70 ms	7

Anatomia opóźnienia obrazu (End-to-End Latency Pipeline)

Opóźnienie obrazu (Latency) w systemach robotycznych i wizyjnych to całkowity czas, jaki upływa od fizycznego zdarzenia w świecie rzeczywistym (np. wejście detalu w pole widzenia) do momentu podjęcia akcji zwrotnej (np. sygnał zatrzymania linii, ruch ramienia robota lub reakcja operatora). Całkowite opóźnienie (End-to-End) składa się z szeregu składowych sprzętowych i programowych:

- **Akwizycja i odczyt sensora (Sensor Readout & Exposure):** Czas potrzebny na ekspozycję matrycy oraz odczyt (readout) ładunków elektrycznych z sensorów. Migawki typu Global Shutter minimalizują zniekształcenia ruchu, ale sam proces odczytu rzędu kilku megapikseli generuje pierwsze opóźnienie (zależne od zegara sensora i interfejsu wewnętrznego, np. SLVS-EC) wynoszące od poniżej 1 ms do kilkunastu milisekund.
- **Transmisja i protokół (Interface & Transport Latency):** Czas potrzebny na spakowanie surowych danych w klatkę i przesłanie jej przez interfejs (np. USB, GigE, CoaXPress). Interfejsy oparte na protokołach bez kompresji (GigE Vision, CXP) cechują się bardzo niskim narzutem czasowym. Standardy z kompresją (H.264/H.265) dodają narzut kodeka związany z analizą ramek, co może wydłużyć proces o 10–150 ms.
- **Buforowanie i sterownik (Host Buffering & OS Driver):** Dane trafiają do bufora karty sieciowej lub frame grabbera, a następnie sterownik systemowy przekazuje je do pamięci RAM za pomocą mechanizmu DMA (Direct Memory Access). Złe zarządzanie kolejką buforów może drastycznie zwiększyć opóźnienie.
- **Przetwarzanie algorytmiczne (Image Processing / AI Inference):** Najbardziej zmienny składnik w systemach automatycznych. Czas przetwarzania przez algorytmy klasyczne (OpenCV) lub sieci neuronowe (detekcja obiektów). Może wynosić od 2 ms (prosty próg binarny) do 50+ ms (głębokie sieci detekcyjne).
- **Wyświetlanie lub Egzekucja (Display / Actuation Latency):** W systemach z operatorem: czas dekodowania strumienia i odświeżania monitora (panele LCD/OLED dodają od 1 do 30 ms). W systemach automatycznych: czas wysłania komendy do sterownika PLC lub kontrolera robota przez magistrale przemysłowe (Profinet, EtherCAT – zazwyczaj poniżej 1–2 ms).

Systemy w pełni automatyczne (Machine Vision & Robotyka)

W pętach sterowania automatycznego (np. systemy 'Robot-in-the-Loop', ramiona typu Delta segregujące produkty, autonomiczne drony), opóźnienie bezpośrednio decyduje o stabilności całego układu fizycznego. Każda milisekunda opóźnienia przesuwają w czasie informację o położeniu obiektu, co przy dużych prędkościach prowadzi do katastrofalnych błędów pozycjonowania.

W systemach automatycznych kluczowe znaczenie ma determinizm czasowy (niski jitter) – system musi gwarantować stałe, powtarzalne opóźnienie. Przykładowo, jeśli ramię robota porusza się z prędkością 2 m/s, to jitter sygnału wizyjnego na poziomie zaledwie 15 ms powoduje, że robot czasem wykonuje manewr chwytania w miejscu oddalonym o 30 mm od rzeczywistego, aktualnego położenia obiektu.

Systemy z człowiekiem-operatorem (Human-in-the-Loop)

Wprowadzenie człowieka do pętli decyzyjnej drastycznie zmienia wymagania oraz zachowanie systemu. Typowe zastosowania to teleoperacja (zdalne sterowanie robotami chirurgicznymi, bezzałogowymi pojazdami wojskowymi, dronami FPV) oraz systemy AR/VR wspomagające pracę na liniach montażowych.

Ludzki mózg i układ nerwowy wykazują specyficzne progi tolerancji na opóźnienia wizualne (glass-to-glass), które decydują o komforcie, precyzji oraz bezpieczeństwie pracy:

- **Poniżej 50 ms (Brak zauważalnego wpływu):** Wartość graniczna w aplikacjach chirurgicznych i teleoperacji precyzyjnej. Opóźnienie poniżej tej wartości jest przez ludzki mózg postrzegane jako natychmiastowe, co pozwala na płynną koordynację wzrokowo-ruchową.
- **50 -150ms (Brak znaczącego wpływu):** Opóźnienie w tym zakresie jest jeszcze przez ludzki mózg postrzegane jako praktycznie natychmiastowe, co pozwala na koordynację wzrokowo-ruchową. Można pisać patrząc w obraz na ekranie.
- **Powyżej 150 ms (Wykrywalne opóźnienie, spadek precyzji):** Operator zaczyna zauważać minimalny dystans czasowy między ruchem własnej ręki a reakcją na ekranie. Spada efektywność pracy, a człowiek zaczyna kompensować opóźnienie poprzez nienaturalne zwalnianie ruchów (strategia 'move-and-wait'). Trudno pisać ręcznie patrząc tylko w ekran.
- **Efekt Motion Sickness (VR/AR):** W systemach nagłownych (VR/AR) opóźnienia z tego przedziału powodują tzw. chorobę symulatorową (Motion Sickness). Dzieje się tak przez konflikt sensoryczny: błędnie rejestruje ruch głowy natychmiast, podczas gdy wzrok widzi przesunięty obraz z opóźnieniem.
- **Powyżej 200 ms (Krytyczne zagrożenie):** Wartość krytyczna dla bezpiecznego sterowania pojazdami lub procesami dynamicznymi. Przekroczenie tego progu uniemożliwia operatorowi natychmiastową reakcję na sytuacje awaryjne i prowadzi do utraty kontroli.

Porównanie rygorów czasowych i progów tolerancji

Typ Systemu / Pętli	Akceptowalne Opóźnienie (End-to-End)	Kluczowy Wymóg Techniczny	Skutki Przekroczenia Progu
Systemy Automatyczne Ultra-Fast (np. Delta robot, sortowanie)	1 - 5 ms	Ekstremalnie niski Jitter (determinizm sprzętowy, trigger z enkodera).	Chybień obiektów, fizyczne uszkodzenia chwytaka, destabilizacja pętli sterowania.
Standardowy Machine Vision (np. kontrola jakości, Pick&Place)	5 - 20 ms	Brak buforowania ramek na hoście, bezpośredni przesył DMA.	Spadek wydajności linii (przepustowości taktowania), opóźnienie reakcji odrzutnika.
Człowiek w petli sterowania	Akceptowalne Opóźnienie (Glass-to-Glass)	Kluczowy Wymóg Techniczny	Skutki Przekroczenia Progu
Krytyczna Teleoperacja z człowiekiem (np. robotyka medyczna, drony)	20 - 50 ms	Wysoki fps (90-120+), brak kompresji wideo lub kodeki ultra-low latency.	Błędy manualne operatora, zmęczenie wzroku, spadek efektywności zabiegu.
Teleoperacja z człowiekiem (np. spawanie prowadzone joystickiem)	100 - 150 ms	fps (25+) brak kompresji wideo,	Błędy manualne operatora, zmęczenie wzroku, spadek efektywności.
Systemy AR/VR (Wspomaganie operatora, gogle nagłowne)	Mniej niż 20 ms (Motion-to-Photon)	Precyzyjny tracking pozycji głowy (IMU), renderowanie z przewidywaniem pozycji.	Nudności, zawroty głowy (Motion Sickness), brak akceptacji technologii przez pracowników.

Latency w systemach z operatorem

Całkowity czas od wystąpienia zdarzenia fizycznego do fizycznej reakcji człowieka (np. wciśnięcia przycisku lub ruchu joystickiem) określa się jako **Total System Latency**.

Składa się ono z dwóch głównych faz: **Glass-to-Glass Latency** (opóźnienie techniczne układu kamera-monitor) oraz **Human Response Time** (opóźnienie biologiczne).

Glass-to-Glass Latency (Opóźnienie Techniczne)

Jest to czas potrzebny na przejście światła przez obiektyw kamery, przetworzenie go na sygnał cyfrowy, przesłanie przez system komputerowy i wyświetlenie jako fotonów na ekranie monitora.

[Zdarzenie] -> (Kamera -> Interfejs -> Komputer -> Monitor) -> [Oko]



Glass-to-Glass Latency

Opóźnienie kamery (Akwizycja i odczyt): 1,5 ms – 35 ms

- **Ekspozycja (Exposure Time):** Czas, przez który matryca zbiera światło. Jeśli kamera pracuje w 60 Hz, maksymalny czas naświetlania jednej klatki to 16,6 ms.
- **Odczyt sensora (Sensor Readout):** Czas potrzebny na przesłanie ładunków z pikseli do przetwornika ADC. W nowoczesnych sensorach (np. Sony Pregius z globalną migawką) wynosi to zazwyczaj **1–4 ms**.
- **Przetwarzanie w kamerze (ISP: Image Signal Processing):** Podstawowy debayering (tworzenie kolorów), korekcja gammy czy balansu bieli realizowane przez procesor kamery (FPGA/ASIC) – dodaje **< 1 ms**.

Interfejs i Transport: 1 ms – 120 ms+

To wąskie gardło zależy od wybranego protokołu transmisyjnego oraz obecności kompresji:

- **Interfejsy bezkompresyjne (GigE Vision, CoaXPress, USB3 Vision, SDI):** Przesyłają surowe ramki bezpośrednio do pamięci komputera (DMA). Opóźnienie transportowe wynosi zaledwie **1–5 ms** (np. dla 10GigE czy 12G-SDI).
- **Interfejsy z kompresją (RTSP, SRT, H.264/H.265, HDMI-over-IP):** Stosowane często w streamingu dalekiego zasięgu. Enkoder w kamerze potrzebuje czasu na przeanalizowanie grupy klatek (GOP). Kompresja i pakietowanie dodają od **15 ms do nawet 100+ ms** opóźnień.

Komputer (Host / Przetwarzanie i System): 2 ms – 40 ms

- **Narzut systemu operacyjnego (OS Buffer & Driver):** Systemy Windows/Linux nie są systemami czasu rzeczywistego (RTOS). Buforowanie pakietów w stosie sieciowym i zarządzanie pamięcią dodaje **2–10 ms**.
- **Przetwarzanie (Application/AI):** Czas, w którym aplikacja pobiera klatkę, nakłada grafikę (OSD/Overlay) lub przetwarza obraz algorytmami. Może to trwać od **2 ms** (proste renderowanie) do **30+ ms** (głębokie sieci neuronowe).

Monitor i Wyświetlanie: 1 ms – 50 ms

- **Przetwarzanie obrazu w monitorze (Input Lag):** Elektronika monitora (skalery, układy poprawy kontrastu, redukcja szumów) potrafi drastycznie opóźnić sygnał. Monitory biurowe i telewizory mają tu od 15 do 50 ms opóźnień. Wyczynowe monitory gamingowe lub medyczne redukują ten czas do **1–5 ms**.
- **Czas odświeżania panelu (Refresh Rate):** Czas oczekiwania na fizyczne narysowanie klatki. Monitor 60 Hz odświeża ekran co 16,6 ms. Monitor 144 Hz skraca ten czas do **6,9 ms**.
- **Czas reakcji piksela (GtG - Gray-to-Gray):** Czas przejścia ciekłego kryształu z jednego stanu w drugi (Smużenie). Dla paneli OLED to **< 0,1 ms**, dla paneli IPS/VA od **1 do 5 ms**.

Human Response Time (Opóźnienie Biologiczne)

Gdy foton z monitora uderza w siatkówkę oka operatora, rozpoczyna się sekwencja procesów neurologicznych i biomechanicznych.

Percepcja Wzrokowa (Visual Processing): ~20 ms – 40 ms

- Światło pobudza fotoreceptory (pręciki i czopki). Sygnał wędruje przez nerw wzrokowy, ciało kolankowate boczne aż do pierwszorzędowej kory wzrokowej (V1) w płacie potylicznym mózgu.

Przetwarzanie Kognitywne i Decyzja: ~50 ms – 150 ms+

- Jest to najbardziej zmienny i złożony komponent zależny od zadania:
 - **Reakcja prosta (Simple RT):** Operator wie, że ma wcisnąć jeden konkretny przycisk, gdy pojawi się jakikolwiek błysk (np. zatrzymanie linii awaryjnej) – trwa ok. **50–80 ms**.
 - **Reakcja z wyborem (Choice RT):** Operator musi ocenić sytuację (np. *czy ten element na taśmie jest uszkodzony w sposób A czy B, i który z dwóch przycisków wcisnąć*) – czas wzrasta do **150 ms – 300 ms+**.

Transmisja Motoryczna (Motor Efferent Pathway): ~10 ms – 20 ms

- Kora motoryczna mózgu generuje impuls ruchu, który wędruje w dół rdzenia kręgowego do odpowiednich grup mięśniowych (np. palca dłoni).

Fizyczna Aktywacja Mięśnia (Muscle Activation & Mechanical Lag): ~30 ms – 70 ms

- Czas potrzebny na pobudzenie włókien mięśniowych (skurcz) oraz fizyczne pokonanie bezwładności kończyny i przesunięcie palca w celu wciśnięcia fizycznego przetącznika lub wychylenia joysticka.