

Wioletta BUJAK

Politechnika Śląska, Instytut Informatyki

## WPLYW WYBORU METOD KODOWANIA OBRAZU I DŹWIĘKU NA JAKOŚĆ WIDEOKONFERENCJI

**Streszczenie.** „Wąskim gardłem” transmisji obrazu i dźwięku w ramach wideokonferencji jest szerokość pasm sieci komputerowej. Przy niemożności szybkiego jej zwiększenia skutecznym narzędziem pozwalającym na polepszenie jakości wideokonferencji okazało się doskonalenie metod kodowania i kompresji wymienionych mediów. W artykule przedstawione zostały wybrane techniki kodowania i kompresji obrazu i dźwięku. Przeanalizowano ich cechy oraz określono, które z nich i w jakim stopniu wpływają na jakość wideokonferencji.

## AN EFFECT OF AUDIO AND VIDEO CODING METHODS CHOICE ON QUALITY OF VIDEOCONFERENCING

**Summary.** A „bottle neck” of audio and video transmission during videoconferencing is a network bandwidth. Because it is impossible to quickly change this situation, the good idea is to improve methods of audio and video coding and compressing to reach better quality of videoconferencing. The purpose of the paper is presenting selected methods of audio and video coding and compressing. After their attributes analysis there is possible to qualify, which of them and how strongly influence on quality of videoconferencing.

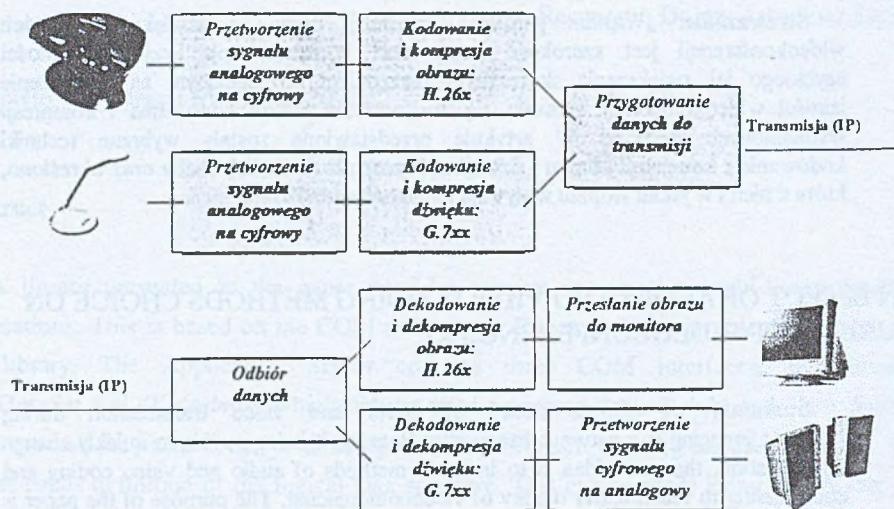
### 1. Jakość wideokonferencji

Przedmiotem niniejszego artykułu jest prezentacja wpływu sposobu kodowania obrazu i dźwięku na jakość wideokonferencji. Analiza jakości wideokonferencji prowadzona była niezależnie dla obu wspomnianych mediów. Po zapoznaniu się ze sposobem funkcjonowania narządów wzroku i słuchu oraz sposobem pobierania i odtwarzania obrazu i dźwięku przez komputerowy system multimedialny możliwe stało się zdefiniowanie pojęcia jakości. Jakość

dźwięku można przedstawić jako funkcję częstotliwości próbkowania dźwięku [kHz], pasma przenoszenia dźwięku, opóźnienia wynikającego z algorytmu kodowania i transmisji przez sieć [ms] oraz stratności algorytmu [%]. Natomiast jakość obrazu może zostać określona przez rozdzielczość obrazu [piksele], ilość rejestrowanych kolorów (tzw. głębię kolorów) [bity], nasycenie barw oraz częstotliwość wyświetlania klatek obrazu [ilość klatek/ sekundę].

## 2. Elementy systemu wideokonferencyjnego

Po zdefiniowaniu jakości wideokonferencji określono elementy systemu wideokonferencyjnego i ustalono, które z nich są znaczące dla dalszych badań.



Rys. 1. Transmisja obrazu i dźwięku w trakcie wideokonferencji  
Fig. 1. An audio and video transmission during videoconferencing

Na jakość dźwięku mają wpływ następujące elementy sprzętowe: mikrofon, jednostka przetwarzająca – obliczeniowa (najczęściej komputer z odpowiednimi kartami rozszerzeń), sieć komputerowa, głośniki / słuchawki, oraz programowe: protokoły kodowania i kompresji dźwięku oraz system operacyjny. Natomiast przy analizie czynników wpływających na jakość obrazu należało zwrócić uwagę na następujące elementy systemu wideokonferencyjnego: kamerę, jednostkę przetwarzającą–obliczeniową (komputer z odpowiednimi kartami



rozszerzeń), sieć komputerową, kartę graficzną, monitor (sprzętowe), zastosowane protokoły kodowania i kompresji obrazu oraz system operacyjny (programowe).

### 3. Elementy decydujące o jakości wideokonferencji

Analiza poszczególnych elementów sprzętowych systemów wideokonferencyjnych pozwoliła na ustalenie, że w dzisiejszych realiach sieć komputerowa jest jednym z najistotniejszych elementów decydujących o jakości prowadzonej wideokonferencji. Przepustowość sieci determinuje ilość możliwych do przesłania danych w jednostce czasu. Ze względu na fakt, iż postęp w dziedzinie zwiększenia przepustowości łączy nie jest tak szybki jak rozwój w dziedzinie parametrów systemów komputerowych, jedynym rozwiązaniem, umożliwiającym transmisję wideokonferencyjną na wystarczająco wysokim poziomie, jest dalszy rozwój technologii kodowania dźwięku i obrazu. Współczesne komputery dysponują wystarczającą mocą oraz pojemnością pamięci operacyjnej, aby znieść koszt coraz to bardziej skomplikowanych technik kodowania. Wybór metody kodowania (tzw. kodeka - ang. Coder/decoder) wpływa na jakość systemu wideokonferencyjnego dwojako: bezpośrednio – zastosowanie określonego algorytmu kodowania i kompresji wpływa na jakość przesyłanego dźwięku i obrazu, oraz pośrednio – stopień kompresji i ilość przesyłanych danych odgrywają znaczną rolę w odniesieniu do zapotrzebowania na stopień zajętości czasu procesora oraz niezbędne pasmo transmisyjne sieci. Całość badań wykonana została przy użyciu dwóch stanowisk wideokonferencyjnych wyposażonych w procesor Pentium 233MHz MMX, 64KB pamięci RAM oraz systemy wideokonferencyjne Intel Business Videoconferencing System [1].

#### 3.1. Kodeki dźwięku

Do testowania wybrano kilka podstawowych kodeków audio wchodzących w skład rekomendacji H.323 (zbioru protokołów umożliwiających prowadzenie wideokonferencji przez sieć pakietową), opisanych w serii G dokumentacji ITU-T [2] oraz technikę kodowania GSM nie należącą do tego standardu. Do testowania kodeków audio zastosowano aplikację *OpenPhone* pozwalającą na ręczny wybór preferowanego kodeka audio, specjalnie dla celów badań przygotowaną aplikację testową *WideoTest* [5] pozwalającą na analizę wielkości zajętego pasma transmisyjnego i przeprowadzenie wideokonferencji oraz aplikację *System Monitor* do analizy stopnia zajętości zasobów systemowych. Jedno ze stanowisk wideokonferencyjnych wykorzystane zostało jako nadawcze, sąsiednie jako odbiorcze. W trakcie badania dla uniknięcia dodatkowych zakłóceń dźwięk transmitowany był jednostronnie przy wyłączonych innych przekazach wideokonferencyjnych.

Jakość dźwięku uzależniona jest w największej mierze od rodzaju wybranego kodeka audio. Poniżej przedstawione zostało zbiorcze porównanie przetestowanych kodeków (szczegółowa analiza kodeków dźwięku: [5]).

Tabela 1

Parametry kodeków audio					
<i>Parametry</i>	<i>G.711a-law</i>	<i>G.711 <math>\mu</math>-law</i>	<i>G.723.1 6,3</i>	<i>GSM 6.1</i>	<i>G.728 LPC</i>
<i>Próbkowanie sygnału</i>	8 kHz	Brak danych	Brak danych	Brak danych	Brak danych
<i>Dane kompresowane</i>	nie	nie	tak	tak	tak
<i>Średnia zajętość pasma</i>	64 Kbps	64 Kbps	10.1 Kbps	6.1 Kbps	18 Kbps
<i>Obciążenie procesora przy nadawaniu[233MHz]</i>	6 %	8 %	8 %	14 %	13 %
<i>Zużycie pamięci nadającego</i>	7 942 KB	6 136 KB	6 228 KB	8 912 KB	9 792 KB
<i>Obciążenie procesora przy odbieraniu[233MHz]</i>	12 %	11 %	11 %	13 %	15 %
<i>Subiektywna ocena jakości</i>	Bardzo dobry	Najlepszy	Dość dobry	Najgorszy	Kiepski

Analizując parametry zebrane w powyższej tabeli oraz mając na uwadze subiektywną jakość otrzymywanego dźwięku, należy się zastanowić nad tym, w jaki sposób można optymalnie dopasować kodek audio, uwzględniając wymagania w stosunku do jakości dźwięku oraz biorąc pod uwagę możliwości systemu videokonferencyjnego. Jak widać, obciążenie systemu komputerowego przy kodowaniu i kompresji dźwięku jest pomijalne.

### 3.1.1. Obciążenie sieci

W realnych warunkach użytkowania sieci pakietowej, gdzie średnie przepustowości mieszczą się w przedziale 3–6 Kbps, nie można sobie pozwolić na zastosowanie takich kodeków jak G.711. Ich stosowanie jest możliwe tylko w sieci ISDN lub w przypadku połączeń w sieci lokalnej (aczkolwiek stosowanie videokonferencji w tym przypadku jest dość rzadko spotykane). Dlatego też najlepszym rozwiązaniem wydaje się być kodek G.723.1, który spośród kodeków z kompresją wyróżnia się bardzo dobrymi parametrami zarówno dotyczącymi jakości głosu, jak i stopnia, w jakim obciąża sieć.

### 3.1.2. Charakterystyka przesyłania danych

Jak można zauważyć, nie zawsze wyższe wymagania kodeka wobec przepustowości pasma świadczą o jego lepszej jakości. Należy też zwrócić uwagę na rozłożenie w czasie danych przesyłanych przez sieć. Charakterystyczny może być tutaj kodek GSM 6.1, który transmituje dane w odstępach 25 milisekundowych, co jest już wykrywane przez ucho ludzkie.



Dodatkowo bardziej równomierne rozłożenie transmisji pozwala na zmniejszenie wielkości bufora zbierającego i porządkującego przesyłane dane i powoduje zminimalizowanie opóźnienie sygnału, co również przyczynia się do zwiększenia jakości dźwięku.

### 3.1.3. Pasma przenoszenia sygnału

Kolejnym aspektem wpływu kodeka na jakość wideokonferencji jest pasmo przenoszenia sygnału. Należy zwrócić tu szczególną uwagę na sam proces kodowania dźwięku wyspecjalizowany w kodowaniu mowy ludzkiej. Transmitowane są jedynie dźwięki odpowiadające jej przedziałowi częstotliwości. Z punktu widzenia wideokonferencji jakość oferowana przez analizowane kodeki jest w większości przypadków całkowicie wystarczająca. Jednakże nie do przyjęcia jest algorytm, w którym zamiast obciążenia, dźwięki o wyższych częstotliwościach zostaną w trakcie procesu kodowania zniekształcone w sposób uciążliwy dla uczestnika, czy też wręcz uniemożliwiający mu kontynuowanie spotkania, jak to zostało zarejestrowane w przypadku kodeka G.728.

## 3.2. Kodeki obrazu

Analiza jakości obrazu została przeprowadzona dla dwóch kodeków: H.263 wchodzącego w skład rekomendacji H.323 i będącego jednocześnie najbardziej rozpowszechnionym kodekiem wideo do prowadzenia wideokonferencji przy użyciu połączeń TCP/IP [3] oraz dla kodeka M-Jpeg, który został dostarczony wraz z aplikacją wideokonferencyjną Cu-SeeMe w wersji 3.12 [4].

Określone zostały parametry przesyłania obrazu przy użyciu kodeka H.263 dla dwóch spośród dostępnych w tym standardzie formatów obrazu: CIF (rozdzielczość 352 x 288 pikseli) oraz QCIF (176 x 144 pikseli). Dla uzyskania jak najbardziej wiarygodnych wyników pomiarów zużycia zasobów procesora i pamięci badania wykonywane były przy ukrytym obrazie lokalnym, braku transmisji obrazu z drugiej strony oraz wyłączonej transmisji głosu. Do badań użyto aplikacji testowej *WideoTest* stworzonej dla celów niniejszych badań oraz aplikacji *System Monitor*.

Poniżej zaprezentowane są uzyskane wyniki charakteryzujące przesyłanie obrazu zakodowanego przy użyciu algorytmu bazującego na specyfikacji H.263 (szczegółowa analiza kodeków obrazu: [5]):

Tabela 2

Parametry kodeka wideo H.263

Parametry	CIF	QCIF
Rozmiar obrazu	352 x 288	176 x 144
Ilość klatek/sekundę	30	30

cd. tabeli 2

<i>Średnia zajętość pasma</i>	77 kbps	80 kbps
<i>Obciążenie procesora przy nadawaniu[233MHz]</i>	16 %	19 %
<i>Obciążenie procesora przy odbieraniu[233MHz]</i>	23 %	26 %
<i>Jakość statycznego obrazu</i>	Bardzo dobra	Dobra

Kodek H.263 [3] jest kodekiem opracowanym optymalnie pod względem wykorzystania w systemach wideokonferencyjnych. Jak można było zaobserwować, w przypadku chwili braku ruchu obraz automatycznie zwiększał swoją dokładność – szczegóły stawały się bardziej wyraźne i zapewniały już nawet przy użyciu formatu QCIF dobrą jakość obrazu. Przy nagłych ruchach można było zaobserwować sposób działania algorytmu zastosowanego w kodeku H.263. Cały obraz dzielony jest na bloki o równych rozmiarach (8 x 8 pikseli) i jest przesyłany różnicowo. Aby uniknąć skomplikowanych obliczeń oraz nie wywoływać zbędnego obciążenia sieci, w pierwszej fazie następuje porównanie obrazu z poprzednią klatką. Jeżeli obraz nie zmienił się, dane nie są przesyłane. Jeżeli natomiast nastąpiła zmiana obrazu, to sprawdzane jest, czy nie było to jedynie przesunięcie obrazu (blok z nowej klatki porównywany jest z sąsiednimi blokami z poprzedniej). Jeśli tak, to wysyłany jest jedynie wektor przesunięcia. Dopiero w sytuacji, gdy nie udało się odnaleźć bloku sąsiedniego, przeprowadzane są przekształcenia niezbędne do zakodowania i skompresowania obrazu. W czasie badań efekt ten został zaobserwowany przy szybszych zmianach, gdy wraz z poruszającym się przedmiotem poruszały się także punkty jego otoczenia. Dodatkowo zaobserwowano, że przy nagłej zmianie obrazu nie jest on od razu kodowany z pełną jakością, ale początkowo przesyłany jest obraz mało dokładny, a dopiero w drugiej kolejności przesyłane są dane potrzebne do prezentacji większej ilości szczegółów.

Do ustalenia parametrów obrazu transmitowanego za pomocą kodeka M-JPEG zastosowana została analogiczna procedura jak w przypadku kodeka H.263. Do generowania obrazu w tym standardzie posłużono się aplikacją *Cu-SeeMe* [4].

Poniżej zaprezentowane są uzyskane wyniki charakteryzujące przesyłanie obrazu:

Tabela 3

Parametry kodeka wideo M-JPEG

<i>Parametry</i>	<i>Minimalna jakość 0%</i>	<i>Jakość 50%</i>	<i>Jakość 75%</i>	<i>Maksymalna jakość 100%</i>
<i>Rozmiar obrazu</i>	352 x 288	352 x 288	352 x 288	352 x 288
<i>Ilość klatek/sekundę</i>	30	1	0.5	0.2
<i>Średnia zajętość pasma</i>	354 kbps	456 kbps	452 kbps	436 kbps
<i>Obciążenie procesora przy nadawaniu[233MHz]</i>	8 %	6 %	6 %	5 %



cd. tabeli 3

<i>Obciążenie procesora przy odbieraniu[233MHz]</i>	7 %	10 %	16 %	5 %
<i>Jakość uzyskanego obrazu</i>	Nie akceptowalna	Dość dobra	Dobra	Dobra

Kodek ten wypadł znacznie gorzej od poprzedniego – jakości statycznej obrazu porównywalnej z H.263 nie uzyskuje się nawet przy wybraniu maksymalnej jakości obrazu, a jest to równoznaczne z wyświetlaniem jednej klatki w pięciosekundowych odstępach (dodatkowo badania wykonane były przy przepustowości dwukrotnie większej niż dla poprzednio rozpatrywanego kodeka). Z wyników badania obciążenia procesora można wnioskować, iż obraz ten jest kodowany stosunkowo prostym algorytmem. Z obserwacji obrazu wynika, że zastosowano w nim elementy mające na celu zmniejszenie ruchu w sieci (przesyłane są te fragmenty, które się zmieniają).

#### 4. Podsumowanie

Powyższa analiza metod kodowania obrazu oraz dźwięku wykazała, jak bardzo jakość wideokonferencji zależy od wybrania odpowiedniego kodeka. Rozwój metod kompresji dźwięku pozwolił nawet na dziesięciokrotne zmniejszenie ilości transmitowanych danych w stosunku do danych nie skompresowanych przy równoczesnej niewielkiej stracie jakości. Również rozwój metod kodowania i kompresji obrazu pozwala na wzrost jakości obrazu przy jednocześnie malejącym wymaganym paśmie transmisji danych. Coraz bardziej zaawansowane techniki kodowania i kompresji obu mediów umożliwiają uzyskanie płynnego i czytelnego obrazu oraz całkowicie zrozumiałego dźwięku. Dlatego też właśnie one dają szansę na rozszerzenie i rozpowszechnienie zastosowań wideokonferencji w sieciach pakietowych.

#### LITERATURA

1. Intel Business Video Conferencing PDK with ProShare Technology.
2. ITU-T Recommendation Series G.
3. ITU-T Recommendation Series H.
4. Cu-SeeMe Documentation.
5. Bujak W.: Praca dyplomowa magisterska: Systemy wideokonferencyjne – badanie i ustalenie parametrów. Promotor prof. A. Grzywak.

Recenzent: Dr inż. Henryk Małysiak

Wpłynęło do Redakcji 12 kwietnia 2002 r.

**Abstract**

The purpose of the paper is presenting how a chosen method of audio and video coding influences a quality of videoconference. The first part of article defines a quality of videoconference and presents elements of videoconference data flow (Fig. 1).

A research of videoconferencing structure and parameters' analyzing brought about conclusion that today a network bandwidth is the most important restriction in videoconferencing.

The method of audio and video coding/decoding influences videoconference in two ways: directly – using different algorithms of coding and compressing data decides about quality of sound and vision; indirectly – the degree of data compression determines if chosen technology allows to perform data transmission in network with bandwidth restriction.

A quality of sound was analyzed for a few codecs (coding/decoding methods) mainly from ITU-T recommendation series G. Table 1 describes the test's results. The best codec should have very good sound quality and use minimal network bandwidth. During testing analyzed was also data flow characteristic and subjective opinion about sound quality.

A quality of vision was analyzed for codec H.263 from ITU-T recommendation series H and M-JPEG distributed with Cu-SeeMe application (ver. 3.12). The results of testing are presented in table 2. The research demonstrates very strongly how a quality of vision depends on chosen codec. Even doubled network bandwidth wasn't enough for M-Jpeg to achieve a quality equal to codec's H.263 video quality.

For testing purposes was prepared special application WideoTest. Additionally was used System Monitor application.

The research has proved that the videoterminal parameters had secondary meaning for quality of videoconferencing. Present computers pass codecs' requirements.

Summarize, the best solution to improve a quality of videoconferencing is to develop coding technology.