



## **KOMPRESJA SYGNAŁÓW AKUSTYCZNYCH PRZY UŻYCIU RÓŻNICOWEJ DYSKRETNEJ TRANSFORMATY KOSINUSOWEJ**

**Acoustic signal compression with the use of differential Discrete Cosine Transform**

**W. Pogribny, M. Drechny**

Wydział Telekomunikacji i Elektrotechniki, Akademia Techniczno-Rolnicza w Bydgoszczy  
[pohry@atr.bydgoszcz.pl](mailto:pohry@atr.bydgoszcz.pl)

### **STRESZCZENIE**

Sygnaly akustyczne analizowane są w dziedzinie czasu i w dziedzinie częstotliwości. Użycie dyskretnej transformaty kosinusowej (DCT) do przetwarzania sygnałów w dziedzinie częstotliwości jest wskazane, ponieważ uzyskujemy większą dokładność przetwarzania niż w dyskretnej transformatce Fouriera (DFT). Ponadto, transformata ta pozwala na wysoką kompresję sygnału, ze względu na nierównomierne widmo DCT skoncentrowane na początkowych jego współczynnikach. W tej pracy zaproponowano uzyskanie widma kosinusowego sygnału, używając niskobitowych algorytmów DCT w formatach delta modulacji (DM). Kompresja widma DCT jest realizowana przez jego modyfikację polegającą na eliminowaniu części współczynników DCT, które są poniżej poziomu szumu kwantowania. Zaproponowane algorytmy są szybkie i pozwalają na przetwarzanie sygnałów w czasie rzeczywistym. W celu zwiększenia szybkości działania procesorów specjalizowanych zaproponowana została metoda użycia takich rodzajów DM, które pozwalają na zamianę operacji mnożenia na operacje przesunięcia bitowego. W związku z tym struktura specjalizowanego procesora staje się prosta.

### **1. WPROWADZENIE**

DCT [1,2] stosowana jest głównie do analizy i kompresji sygnałów bazując na swoich własnościach, którymi są: koncentracja widma sygnału na jego początkowych współczynnikach, użycie tych samych funkcji bazowych do obliczenia DCT oraz odwrotnej DCT (IDCT), obliczenia wykonywane w dziedzinie liczb rzeczywistych i możliwość obliczeń DCT za pomocą algorytmów DFT. Podstawowy algorytm DCT [1,2] oraz wszystkie szybkie odmiany tego algorytmu obliczane są w wielobitowym formacie PCM. Format ten zapewnia szerokie pasmo przetwarzania oraz dużą dokładność obliczeń. Jednak algorytmy DCT w formacie PCM zwierają czasochłonne operacje mnożenia. Eliminując je można zwiększyć szybkość obliczeń. Jedną z metod jest zastosowanie niskobitowego zmodyfikowanego formatu DPCM (MDPCM) w którym kroki kwantowania są proporcjonalne do potęgi liczby 2 [3]. Operacje mnożenia zastępowane są szybkimi operacjami przesunięcia bitowego SHIFT. Ponadto widmo sygnału przetwarzanego

w formacie MDPCM lub innych formatach różnicowych można modyfikować za pomocą częściowego zerowania jego współczynników przez co zwiększa się jakość odtwarzanych sygnałów po IDCT. Innym sposobem usunięcia operacji mnożenia jest zastosowanie syntetycznego kodu DPCM złożonego z kilku połączonych równolegle koderów MDPCM.

Celem niniejszej pracy jest opracowanie i zbadanie szybkich i dokładnych algorytmów DCT w formatach różnicowych, które nie zawierają operacji mnożenia a także pozwalają na istotną kompresję danych.

## 2. SZYBKI ALGORYTM DCT DLA FORMATÓW RÓŻNICOWYCH

Zastępując format PCM w podstawowym algorytmie DCT [1,2] różnicami wzrastającymi  $\nabla$ PCM a następnie transformując sumy przyrostów sygnału oraz sumy kosinusów na płaszczyźnie tak jak to zostało pokazane w [3] i [4], otrzymujemy szybki algorytm DCT w formacie różnicowym. Zamieniając różnice na kroki kwantowania DM gdy nie występują błędy przeciążenia kodera, otrzymujemy szybki algorytm DCT przedstawiony poniżej:

$$X(m) = k_m \sum_{n=0}^{N-1} s_n \sum_{r=n}^{N-1} \cos\left(\frac{\pi m(2r+1)}{2N}\right) \quad (1)$$

gdzie:

$N$  – ilość próbek wyjściowych oraz ilość współczynników DCT,  $m, n = \overline{0, N-1}$  i

$$k_m = \begin{cases} \frac{\sqrt{2}}{N}, & \text{for } m = 0 \\ \frac{2}{N}, & \text{for } m = \overline{1, N-1} \end{cases}.$$

Kroki kwantowania  $s_n$  z algorytmu (1) są krokami kwantowania koderów DM takich jak DPCM, ADM czy LDM. Modyfikując kroki kwantowania kodera DPCM tak aby były proporcjonalne do potęgi liczby 2 ( $2^x$ , gdzie  $x \in \mathbb{N}$ ), możemy wyeliminować z algorytmu (1) czasochłonne operacje mnożenia i zastąpić je szybkimi operacjami przesunięcia bitowego SHIFT. Ponadto wcześniejsze obliczenie i zapamiętanie sum kosinusów w pamięci procesora specjalizowanego zmniejsza ilość operacji z  $N^3$  na  $N^2$  wykonywanych z niską bitowością DM i umożliwia pracę algorytmu w systemach czasu rzeczywistego.

W celu zwiększenia jakości odtwarzania sygnałów po transformacji DCT w formacie różnicowym, zastosowano zerowanie części współczynników widma DCT czyli przypisywanie obliczonym wartościom współczynników DCT wartości liczbowej równej 0. Operację zerowania można przeprowadzić wieloma metodami. W niniejszej pracy przedstawiono dwie metody. Pierwsza, w której początkowym współczynnikiem zerowania jest współczynnik określający częstotliwość górną sygnału. Natomiast końcem przedziału zerowania jest ostatni współczynnik DCT przetwarzanego sygnału. Druga metoda polega na założeniu liczby następujących po sobie współczynników oraz wartości poziomów granicznego współczynnika, poniżej którego wartości poziomów współczynników są traktowane jako szum kwantowania. Rozpoczynając od pierwszego współczynnika DCT szukamy pierwszej grupy następujących po sobie współczynników o założonej ilości,

których poziomy co do modułu są mniejsze od założonej wartości granicznej poziomu. Pierwszym elementem od którego rozpoczyna się zerowanie jest ostatni element ze znalezionej grupy współczynników, a koniec zerowania wyznaczany jest przez ostatni współczynnik DCT tego sygnału.

Metoda tworzenia syntetycznego kodu DPCM (SDPCM) polega na tworzeniu wszystkich kroków kwantowania kodera DPCM za pomocą sumowania odpowiednich kroków koderów MDPCM, których wartości są proporcjonalne do potęgi liczby 2. Przykładowa realizacja 4 bitowego kodu SDPCM przedstawia tabela 1.

Tabela 1. Przykład realizacji 4 bitowego kodu SDPCM

Krok DPCM	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
MDPCM #1	1	2	-4	-4	-2	-2	-1	0	1	2	2	4	4	4	-1
MDPCM #2	-8	-8	-1	0	-1	0	0	0	0	0	1	0	1	2	8

Kody MDPCM#1 oraz MDPCM#2 są kodami 3 bitowymi o krokach kwantowania wybieranych ze zbioru: MDPCM#1 {0, ±1, ±2, ±4}, MDPCM#2 {0, ±1, 2, ±8}. W podobny sposób można tworzyć kody SDPCM dla słów kodowych o długościach np. 5, 6, 7 czy 8 bitów. Podczas obliczeń z użyciem zaproponowanego algorytmu DCT, kroki kwantowania MDPCM sterują przesunięciami bitowymi sum kosinusów ze wzoru (1).

### 3. WYNIKI BADAŃ

Dokładność zaproponowanych algorytmów została zbadana przetwarzając sygnały mowy, którymi było kilka polskich słów. Sygnały mowy przetwarzano dla 2 różnych częstotliwości próbkowania, dla długości słów kodowych 3, 4, 5 i 6 bitów w formatach PCM, DPCM (SDPCM), MDPCM oraz ADM. Badania powtórzono dla tych samych sygnałów mowy przetwarzanych przy założeniu takich samych parametrów (częstotliwość próbkowania, długość słów kodowych oraz formaty) lecz z zastosowaniem zerowania części współczynników widma DCT. Podczas badań zastosowano pierwszą metodę zerowania opisaną w punkcie 2. Tabela 2 przedstawia wyniki badań sygnałów mowy bez zerowania natomiast tabela 3 z zerowaniem.

Najlepsze wyniki bez zerowania części współczynników DCT dla bitowości 3 oraz 4, otrzymujemy dla sygnałów przetwarzanych w formacie MDPCM. Powyżej 4 bitów większą dokładność przetwarzania zapewnia algorytm DCT w formacie DPCM i SDPCM. Dla algorytmu PCM o niskiej bitowości oraz jednobitowego ADM otrzymujemy najmniejsze wartości Signal-to-Noise Ratio (SNR).

Tabela 2. Uśrednione SNR dla badanych sygnałów mowy bez zerowania

Liczba bitów	Częstotliwość próbkowania: 11025Hz				Częstotliwość próbkowania: 22050Hz			
	PCM	DPCM SDPCM	MDPCM	ADM	PCM	DPCM SDPCM	MDPCM	ADM
	SNR [dB]							
3	11,14	19,26	20,81	10,90	11,19	23,94	25,78	16,25
4	15,68	25,03	26,90		15,67	29,93	32,83	
5	20,76	30,50	26,95		20,69	35,81	32,88	
6	26,67	36,16	-		26,65	41,55	-	

Tabela 3. Uśrednione SNR dla badanych sygnałów mowy z zerowaniem

Liczba bitów	Częstotliwość próbkowania: 11025Hz				Częstotliwość próbkowania: 22050Hz			
	PCM	DPCM SDPCM	MDPCM	ADM	PCM	DPCM SDPCM	MDPCM	ADM
	SNR [dB]							
3	11,75	20,18	22,03	11,22	12,08	25,81	27,84	19,63
4	16,43	26,13	28,73		16,90	31,83	36,34	
5	21,62	31,64	28,77		22,23	36,95	36,36	
6	27,71	36,77	-		28,41	40,39	-	

Zastosowanie prostej operacji zerowania części widma DCT spowodowało wzrost wartości SNR w każdym badanym przypadku. Największy przyrost SNR zaobserwować można dla formatu MDPCM.

Analizując tabelę 2 i 3 można zauważyć, że największą wartość SNR otrzymujemy dla zerowanych formatów MDPCM dla bitowości 3 i 4 oraz zerowanego formatu DPCM i SDPCM dla bitowości 5 i 6 niezależnie od częstotliwości próbkowania.

## 5. WNIOSKI

Zaproponowane różnicowe algorytmy DCT są szybkie ponieważ wyeliminowano z nich czasochłonne operacje mnożenia i zastąpiono je szybkimi operacjami przesunięcia kodu SHIFT. Taka zamiana jest możliwa przy użyciu zmodyfikowanej DPCM w której kroki kwantowania są proporcjonalne do potęgi liczby 2. Zastosowanie kilku równoległe działających i uzupełniających się koderów MDPCM umożliwi działanie w formacie SDPCM, który łączy dużą szybkość działania (brak operacji mnożenia) i wysoką dokładność przetwarzania (kod DPCM).

Kompresję widma DCT sygnału przetwarzanego w formatach różnicowych badanych w tej pracy dla słów kodowych 3 i 4 bitowych należy zrealizować za pomocą zerowania części współczynników jego widma co nie pogarsza jakości odtwarzanego.

Procesory zbudowane w oparciu o zaproponowane algorytmy są szybkie a ich struktura jest prosta przez co mogą być stosowane w systemach czasu rzeczywistego.

## LITERATURA

1. N. AHMED, T. NATARJAN, K. R. RAO, Discrete Cosine Transform, IEEE Transactions on Computers, vol. 23, strony 90-93 (Styczeń 1974).
2. K. R. RAO, P. YIP, Discrete cosine transform: algorithms, advantages and applications, Academic Press, 1990.
3. W. POGRIBNY, M. DRECHNY, Szybka transformata DCT sygnału mowy z wykorzystaniem formatów różnicowych, Krajowe Sympozjum Telekomunikacji, Bydgoszcz 2003, strony 9-16.
4. W. POGRIBNY, M. DRECHNY, Fast Algorithms of Discrete Fourier and Cosine Transforms in Mixed Formats, 3<sup>th</sup> IEEE International Symposium on Image and Signal Processing and Analysis ISPA'03, Rzym 2003, strony 779-783.