

Sveučilište u Zagrebu
Fakultet elektrotehnike i računarstva

SEMINARSKI RAD

iz kolegija: Napredne metode digitalne obrade signala

NAPREDNE METODE KOMPRESIJE ZVUČNIH SIGNALA

Anita Cazin	0036400516
Mirjana Cazin	0036377333

Zagreb, siječanj 2007.

SADRŽAJ

UVOD	3
MDCT	4
MDCT POMOĆU DCT-IV I GIVENS-OVIH ROTACIJA	6
DCT-IV POMOĆU GIVENS-OVIH ROTACIJA	9
STRUKTURA PODIZANJA	9
ZAKLJUČAK	13
LITERATURA	14

UVOD

Razvojem elektroničkih uređaja i povećanjem kapaciteta memorija, na prvom mjestu kod kodiranja audio signala više nije što bolja kompresija uz gubitke, već kvaliteta dekodiranog signala, odnosno kompresija bez gubitaka. Većina danas poznatih transformacija kao izlazne vrijednosti daje decimalne vrijednosti, te je nužna kvantizacija tih vrijednosti. Kvantizacija unosi pogreške i nije više moguća potpuna rekonstrukcija takvog signala. Ljestvičastim strukturama koje omogućavaju primjenu nelinearnih i vremenski promijenjivih stupnjeva transformacije, uz potpunu rekonstrukciju, ostvarena je IntMDCT transformacija, koja uz zadržavanje dobrih svojstava MDCT transformacije daje cjelobrojne izlazne vrijednosti. U ovom seminarskom radu opisano je dobivanje MDCT transformacije pomoću DCT-IV transformacije te razlaganje DCT-IV transformacije na Givens-ove rotacije. Daljnjim razlaganjem Givens-ovih rotacija izvedena je ljestvičasta struktura. U stupnjevima podizanja ljestvičaste strukture izvršeno je zaokruživanje vrijednosti, te je tako dobivena IntMDCT transformacija.

MDCT

Većina aktualnih audio kodiranja temelji se na transformacijama kao što je Modificirana Diskretna Kosinusna Transformacija - MDCT (Modified Discrete Cosine Transform). MDCT je transformacija temeljena na DCT-IV, sa dodatnim svojstvom „preklapanja“, odnosno napravljena je tako da se obavlja na uzastopnim blokovima podataka koji se preklapaju na taj način da druga polovica jednog bloka prekrije prvu polovicu sljedećeg bloka. Ovakvo preklapanje, uz kvalitete sažimanja energije DCT transformacije, čini MDCT idealnim za primjene u kompresiji signala, jer pomaže da se izbjegnu artefakti koji se javljaju na granicama blokova. Također uslijed preklapanja blokova postizemo poništavanje aliasinga u vremenskoj domeni (TDAC - Time Domain Aliasing Cancellation) koje se javlja uslijed decimacije u frekvencijskoj domeni. Navedene transformacije za ulazne cjelobrojne vrijednosti daju izlazne vrijednosti u obliku decimalnog broja, te je nužan proces kvantizacije odnosno zaokruživanja.

Blok signala $x_t(k)$, sa $2N$ uzoraka u vremenskoj domeni ($k=0, \dots, 2N-1$) upotrebljen je za izračunavanje N linija spektra $X_t(m)$, $m=0, \dots, N-1$. Dva bloka koji slijede jedan iza drugoga preklapaju se za 50 % svoje duljine tako da svaki blok proizvede N novih uzoraka u vremenskoj domeni. Za preklapanje blokova bez teškoća upotrebljava se prozor $w(k)$, gdje je $k=0, \dots, N-1$.

Formula za MDCT transformaciju :

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} \omega(k) x_t(k) \cos\left(\frac{\pi}{4N} (2k+1+N)(2m+1)\right), \quad (1)$$
$$m=0, \dots, N-1$$

inverzna MDCT :

$$y_t(k) = \omega(k) \sqrt{\frac{2}{N}} \sum_{m=0}^{N-1} x_t(m) \cos\left(\frac{\pi}{4N} (2k+1+N)(2m+1)\right), \quad (2)$$
$$k=0, \dots, 2N-1$$

Primjenom MDCT i inverzne MDCT transformacije unosi se pogreška radi aliasinga u vremenskoj domeni. Ta pogreška može se poništiti zbrajanjem izlaza inverzne MDCT transformacije dva susjedna bloka t i $t+1$ u preklapajućem dijelu:

$$x'_t(k) = y_t(N+k) + y_{t+1}(k), \quad (3)$$

$$k=0,\dots,N-1$$

Prozori koje koristimo za preklapanje blokova, moraju ispuniti određene uvjete, da bi se osiguralo poništenje aliasinga u vremenskoj domeni. Dovoljan uvjet za poništenje aliasinga u vremenskoj domeni je:

$$\omega(k)^2 + \omega(N+k)^2 = 1 \quad (4)$$

$$\omega(k) = \omega(2N-1-k) \quad (5)$$

$$k = 0, \dots, N-1.$$

Primjer prozora koji ispunjava ove uvjete je sinusni prozor:

$$\omega(k) = \sin\left(\frac{\pi}{4N}(2k+1)\right), \quad (6)$$

$$k=0,\dots,2N-1$$

MDCT POMOĆU DCT-IV I GIVENS-OVIH ROTACIJA

MDCT transformacija sa prozorom duljine $2N$ se može izračunati DCT transformacijom (tip 4) duljine N . To se postiže izvođenjem preklapanja u vremenskoj domeni (TDA – Time Domain Aliasing) i naknadnom primjenom DCT-a (diskretne kosinusne transformacije).

Definirajmo preklapljeni signal u vremenskoj domeni $\tilde{x}_t(k)$, $k=0, \dots, N-1$:

$$\tilde{x}_t(k) = \omega\left(\frac{N}{2} + k\right)x_t\left(\frac{N}{2} + k\right) - \omega\left(\frac{N}{2} - 1 - k\right)x_t\left(\frac{N}{2} - 1 - k\right) \quad (7)$$

$$\tilde{x}_t(N-1-k) = \omega\left(\frac{3N}{2} + k\right)x_t\left(\frac{3N}{2} + k\right) + \omega\left(\frac{3N}{2} - 1 - k\right)x_t\left(\frac{3N}{2} - 1 - k\right), \quad (8)$$

$$k=0, \dots, N/2 - 1.$$

Tada se formula za MDCT svodi na:

$$X_t(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} \tilde{x}_t(N-1-k) \cos\left(\frac{\pi}{4N} (2k+1)(2m+1)\right), \quad (9)$$

$$m=0, \dots, N-1$$

što je upravo jednako primjeni DCT transformacije duljine N na signal $-\tilde{x}_t(N-1-k)$, $k=0, \dots, N-1$.

Lijeva polovica prozora za blok t se preklapa sa desnom polovicom za blok $t-1$. Iz jednadžbe (8) slijedi da je taj dio ulaznog signala upotrebljen za MDCT transformaciju bloka $t-1$:

$$\begin{aligned} \tilde{x}_{t-1}(N-1-k) &= \omega\left(\frac{3N}{2} + k\right)x_t\left(\frac{N}{2} + k\right) + \omega\left(\frac{3N}{2} - 1 - k\right)x_t\left(\frac{N}{2} - 1 - k\right) \\ \tilde{x}_{t-1}(N-1-k) &= \omega\left(\frac{N}{2} - 1 - k\right)x_t\left(\frac{N}{2} + k\right) + \omega\left(\frac{N}{2} + k\right)x_t\left(\frac{N}{2} - 1 - k\right) \end{aligned}, \quad (10)$$

$$k=0, \dots, N/2 - 1.$$

Kombinacijom gornje jednadžbe sa jednadžbom (8) za blok t , vidi se da je u preklapajućim dijelovima dva susjedna bloka $t-1$ i t , signal u vremenskoj domeni $x_t(k)$, $K=0, \dots, N-1$ pripremljen za primjenu DCT-a (diskretne kosinusne transformacije) sa:

$$\begin{pmatrix} \tilde{x}_t(k) \\ \tilde{x}_{t-1}(N-1-k) \end{pmatrix} = \begin{pmatrix} \omega(\frac{N}{2}+k) & -\omega(\frac{N}{2}-1-k) \\ \omega(\frac{N}{2}-1-k) & \omega(\frac{N}{2}+k) \end{pmatrix} \begin{pmatrix} x_t(\frac{N}{2}+k) \\ x_t(\frac{N}{2}-1-k) \end{pmatrix}, \quad (12)$$

$k=0,\dots,N/2-1.$

Iz uvjeta za TDAC u jednadžbi (4) slijedi:

$$\omega(\frac{N}{2}+k)^2 + \omega(\frac{N}{2}-1-k)^2 = 1 \quad (13)$$

tako da za određene kutove definirane formulom:

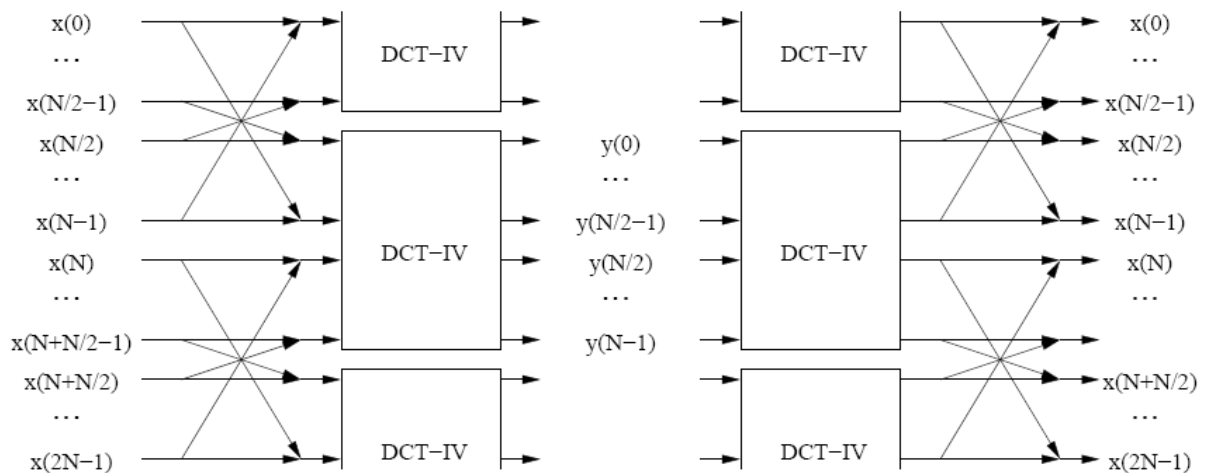
$$\alpha_k = \arctan \frac{\omega(\frac{N}{2}-1-k)}{\omega(\frac{N}{2}+k)}, \quad (14)$$

$$k=0,\dots,N/2-1$$

ta predobrada u vremenskoj domeni može biti zapisana primjenom Givens-ovih rotacija

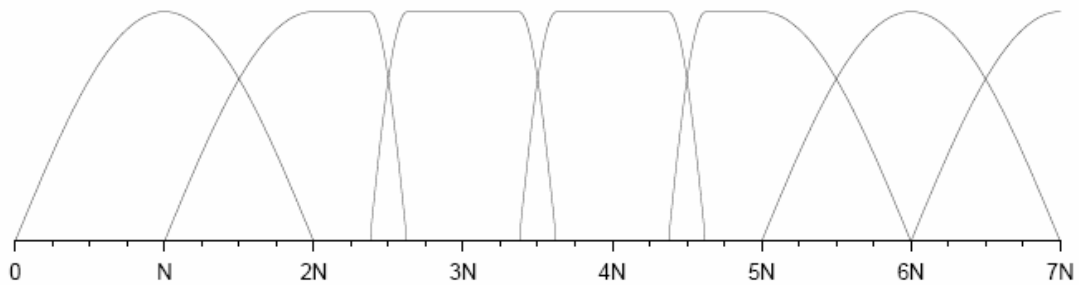
$$\begin{pmatrix} \cos \alpha_k & -\sin \alpha_k \\ \sin \alpha_k & \cos \alpha_k \end{pmatrix}. \quad (15)$$

Za inverznu MDCT isti postupak može biti upotrebljen ali obrnutim redoslijedom. Inverzna DCT-IV je zapravo DCT-IV. Inverz rotacija primjenjenih za prozore i aliasing u vremenskoj domeni postiže se primjenom rotacija sa kutovima $-\alpha_k, k=0,\dots,N/2-1.$ Cijeli proces je prikazan slikom 1.



Slika 1. Dekompozicija MDCT i inverzne MDCT transformacije u Givens-ove rotacije i DCT-IV transformaciju

Ovakvom dekompozicijom MDCT transformacije lako se može uočiti da oblik prozora može biti odabran individualno u svakom bloku. Temeljeno na rotacijama, prilagodba prozora može biti izvedena promjenom kuta rotacije. Za potpunu rekonstrukciju dovoljno je samo odabrati negativne vrijednosti kutova u svakom bloku prilikom inverzne transformacije. Jedan od mogućih poredaka prozora prikazan je slikom 2.



Slika 2. Tipičan slijed oblika prozora za MDCT transformaciju

DCT-IV POMOĆU GIVENS-OVIH ROTACIJA

Diskretna kosinusna transformacija tipa IV (DCT-IV) duljine N definirana je:

$$X_i(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} x(k) \cos\left(\frac{\pi}{4N} (2k+1)(2m+1)\right), \quad (16)$$

$$m=0, \dots, N-1$$

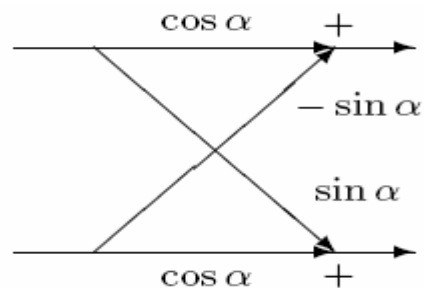
Koeficijenti DCT-IV čine ortonormalnu matricu dimenzija $N \times N$. Svaka ortonormalna $N \times N$ matrica može se rastaviti na $\frac{N(N-1)}{2}$ Givens-ovih rotacija, ali ta dekompozicija nije jedinstvena, pa su moguće i dekompozicije koje koriste što je moguće manji broj rotacija. Brzi algoritmi za DCT-IV temelje se na smanjivanju broja rotacija na $O(N \log_2 N)$. U daljnjem tekstu biti će opisana jedna od mogućih dekompozicija DCT-IV kao brzi algoritam za MDCT.

STRUKTURA PODIZANJA

Primjena Givens-ove rotacije

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

je prikazana slikom 3.

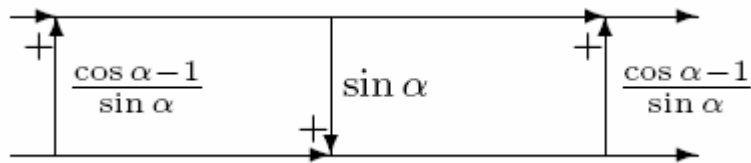


Slika 3. Givens-ova rotacija

Givens-ove rotacije se mogu rastaviti u tri koraka podizanja:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \quad (17)$$

Dekompozicija je prikazana slikom 4.



Slika 4. Givens-ova rotacija u tri koraka podizanja

Sad možemo uključiti funkciju zaokruživanja

$$r: \mathbb{R} \rightarrow \mathbb{Z}$$

u svaki od koraka podizanja da bi dobili cjelobrojnu aproksimaciju. Drugi korak podizanja

$$(x_1, x_2) \rightarrow (x_1, x_2 + x_1 \sin \alpha)$$

možemo aproksimirati sa

$$(x_1, x_2) \rightarrow (x_1, x_2 + r(x_1 \sin \alpha)).$$

U ovom slučaju prva komponenta nije promijenjena, stoga se $r(x_1 \sin \alpha)$ može izračunati i nakon primjene koraka podizanja. Inverzni korak definiramo:

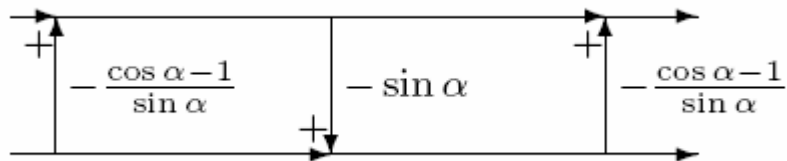
$$(x'_1, x'_2) \rightarrow (x'_1, x'_2 + r(x'_1 \sin \alpha)).$$

Stoga cjelobrojna aproksimacija koraka podizanja može biti rekonstruirana bez uvođenja bilo kakve greške. Primjenom ovakve aproksimacije na svaki od tri koraka podizanja može se dobiti cjelobrojna aproksimacija Givens-ove rotacije. Ta cjelobrojna rotacija može biti obrnuta bez uvođenja grešaka, primjenom inverznih koraka zaokruživanja u obrnutom redosljedu, upotrebom istih funkcija zaokruživanja.

Ukoliko je funkcija zaokruživanja r neparno simetrična, inverzna zaokružena rotacija je identična zaokruženoj rotaciji sa kutom od $-\alpha$:

$$\begin{pmatrix} \cos \alpha_k & \sin \alpha_k \\ -\sin \alpha_k & \cos \alpha_k \end{pmatrix}. \quad (18)$$

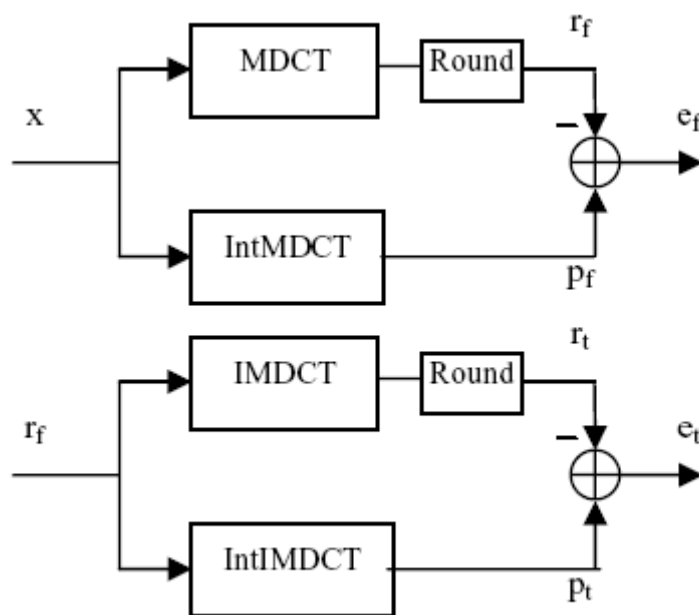
Na slici 5 prikazana je inverzna rotacija pomoću koraka podizanja.



Slika 5. Inverzna Givens-ova rotacija u tri koraka podizanja

Zamjenom pojedine Givens-ove rotacije MDCT dekompozicije cijelobrojnim rotacijama, izlazne vrijednosti ostaju cijelobrojne, za cijelobrojne ulazne vrijednosti. Cijeli proces je reverzibilan primjenom inverzne rotacije obrnutim redosljedom. Tako je dobivena cijelobrojna aproksimacija MDCT transformacije kojom je očuvana potpuna rekonstrukcija.

U [2] predstavljena je jedna od mogućih dekompozicija Givensovih rotacija za implementaciju IntMDCT transformacije. Pokazalo se je da je ukupan broj operacija zaokruživanja u tom algoritmu $2.5N$, što rezultira mnogo manjom greškom aproksimacije. Također se može procijeniti da je kompleksnost operacija računanja DTC-a sa cijelobrojnim vrijednostima otprilike upola manja nego računanja DTC-a sa vrijednostima sa pomičnim zarezom. Predloženi algoritam uspoređen je sa algoritmom koji se koristi za MPEG4 kodek bez gubitaka nazvan Reference Model 0 (RM0). Slika 6 pokazuje blok dijagram za procjenu performansi transformacija.



Slika 6. Blok dijagram za procjenu performansi transformacija

Da bi usporedba bila moguća, preuzeti su uvjeti procjene kvalitete kodiranja kod MPEG4 direktne realizacije RM0 (450 sekundi, 15 različitih vrsta muzike, na 48kHz/16bita i 96kHz/24bita) i izračunata srednja kvadratna pogreška (MSE – Mean Square Error) prema izrazu:

$$MSE = \frac{1}{L} \sum_{j=0}^{L-1} \frac{1}{1024} \sum_{i=0}^{1023} e_i^2 . \quad (19)$$

U tablici 1 prikazani su rezultati usporedbe.

	RM0 (Direct Implementation)	Proposed Algorithm	Compression ratio improvement
(IntMDCT) *MSE	1.789	0.547	0.87%
(IntIMDCT) *MSE	1.774	0.545	
(IntMDCT) **MSE	1.785	0.546	0.42%
(IntIMDCT) **MSE	1.766	0.546	
Complexity (instructions)	47,616	34,816	

*48kHz/16 bit test set; **96kHz/24 bit test set

Tablica 1. Rezultati usporedbe

Vidljivo je da su postignuti značajno bolji rezultati kompresije, bez ikakvih novih operacija, a ujedno je i kompleksnost operacija smanjena za otprilike 26%.

ZAKLJUČAK

IntMDCT transformacija za ulazne cjelobrojne vrijednosti ima izlazne cjelobrojne vrijednosti. Pruža i potpunu rekonstrukciju, tako da nije unesena ni jedna greška primjenom transformacije i inverzne transformacije. Izvedena je iz Modificirane Diskretne Kosinusne Transformacije (MDCT), te su stoga očuvane njene prednosti: ima preklapajuće blokove, što pruža bolju frekvencijsku selektivnost od transformacija koje ne rade sa preklapajućim blokovima, zbog poništavanja aliasinga u vremenskoj domeni (TDAC) održana je frekvencija graničnog uzorkovanja, tako da je ukupan broj vrijednosti spektra koje predstavljaju audio signal ne premašuju broj ulaznih uzoraka.

LITERATURA

- [1] R. Geiger, T. Sporer, J. Koller, K. Brandenburg: „Audio Coding based on Integer Transforms“, Fraunhofer Institut für Integrierte Schaltungen, Arbeitsgruppe für Elektronische Medientechnologie i Ilmenau Technical University, Ilmenau, Njemačka, 2001
- [2] H. HaiBin, R. Susanto, Y. Rongshan, L. Xiao: „A Fast Algorithm of Integer MDCT for Lossless Audio Coding“, Agency for Science, Technology and Research (A*STAR), Institute for Infocomm Research (I2R), Singapore, 2004
- [3] D. Seršić: Predavanja iz kolegija Napredne metode digitalne obrade signala
- [4] http://en.wikipedia.org/wiki/Discrete_cosine_transform
- [5] http://en.wikipedia.org/wiki/Modified_discrete_cosine_transform