

Systemy multimedialne Kodowanie dźwięków.

Paweł Szoltysek

Spis treści

1	Wstęp	1
2	Analiza i omówienie formatów i narzędzi bezstratnego i stratnego kodowania dźwięków	2
2.1	Formaty zapisu danych dźwiękowych	2
2.1.1	Zapis muzyki na nośnikach	2
2.1.2	Standardowe formaty muzyczne	4
2.1.3	Kontenery multimedialne	6
2.1.4	Pliki typu Musical Instrument Digital Interface (MIDI)	8
2.2	Aplikacje do kodowania plików audio	8
3	Ilustracja cech wybranych metod kodowania	9
3.1	Kodowanie muzyki pop	9
3.2	Kodowanie muzyki klasycznej	10
3.3	Kodowanie mowy	11
4	Podsumowanie	11

1 Wstęp

Tworzenie i edycja muzyki jest dziś naturalnym obszarem zastosowania komputerów. Przez ostatnie kilkanaście, a nawet kilkadziesiąt lat komputery wywierały ogromny wpływ na tworzoną muzykę (bez nich nie powstałyby takie gatunki muzyki jak na przykład techno), a także na video dołączane do muzyki, które w dzisiejszych czasach posiada zdecydowanie wyższy poziom jakościowy.

Prosta komputerowa obróbka stosowana jest na przykład przez radio Zet, które od wielu lat nadaje krótkie audycje takie jak: "Rozmowy rolowane", gdzie na zadane przez reportera pytanie konstruowana jest odpowiedź na podstawie wcześniejszych wypowiedzi rozmówcy (po prostu zostaje skopiowany nagrany wcześniej fragment ścieżki dźwiękowej), zwykle w sposób zabawny i prześmiewczy.

W niniejszej pracy skupię się na różnych metodach zapisu oraz możliwościach kodowania plików audio. W historii plików audio, przełomowym momentem było rozpowszechnienie się formatu MP3 do użytku domowego. W czasach, kiedy dyski miały conajwyżej kilka gigabajtów pojemności, a procesory po kilkaset megahertz, pozwalał on w czasie rzeczywistym dekodować wysokiej jakości muzykę zapisaną w niewielkim, kilkumegabajtowym pliku.

Od tego czasu format ten nadal rządzi niepodzielnie w zastosowaniach domowych. Jednak obecny wysokiej klasy sprzęt hi-fi nierzadko odtworza taką muzykę niepoprawnie (z niewystarczającą jakością), co sprawia, że pliki takie koduje się z wyższym wskaźnikiem bitrate. Obecnie trwają też prace nad przedstawieniem formatu muzyki w wysokiej jakości, który przyjąłby się z podobnym sukcesem do MP3.

2 Analiza i omówienie formatów i narzędzi bezstratnego i stratnego kodowania dźwięków

2.1 Formaty zapisu danych dźwiękowych

Lata rozwoju zaowocowały pojawieniem się wielu różnych standardów zapisu komputerowego muzyki. Poniżej zostaną przedstawione główne formaty, które są dziś popularne.

2.1.1 Zapis muzyki na nośnikach

Zacznę od opisu standardów zapisu plików audio na nośnikach typu CD-Audio, SACD oraz DVD-Audio.

Compact Disc Digital Audio (CD-Audio, CD-DA)

Standard CD-Audio jest domyślnym formatem zapisu muzyki na płytach kompaktowych i istnieje od momentu ich powstania. Płyty CD stworzono właśnie w celu przenoszenia dźwięku. Dekodowanie ich nie wymaga dużej mocy obliczeniowej, co spowodowało rozpowszechnienie się tego nośnika bezpośrednio po jego zaprezentowaniu w roku 1982.

Format ten wykorzystuje kodowanie PCM o częstotliwości próbkowania 44,1 kHz i rozdzielczości 16 bitów na próbkę. Wartości te są dobrane nieprzypadkowo. Częstotliwość próbkowania dźwięku tak zapisanego pozwala (wg twierdzenia Nyquista-Kotelnikova-Shannona) na wierne odwzorowanie dźwięku o częstotliwości do 22,05 KHz. Powinna być ona wystarczająca, przy założeniu faktu, że człowiek słyszy sygnały o częstotliwości z zakresu 20Hz - 20kHz. Rozdzielczość 16-bitowa pozwala na zapisanie wysokości sygnału od wartości 0 do 32768, co daje nam dynamikę rzędu 90dB ($20 \times \log 32768 = 90, 3$). Oznacza to, że najgłośniejszy sygnał ma energię około miliard razy większą od najcichszego. Jest to zdecydowany atut tego nośnika względem jego

konkurentów - taśmy magnetofonowej i płyty analogowej. Są to jednak wyliczenia czysto analityczne - duży wpływ na jakość ma sposób nagrania, oraz sprzęt odtwarzający.

Sygnal w ten sposób kodowany i przechowywany jest stereofoniczny, w systemie 2.0.

Bitrate tego formatu wynosi $44110 \times 16 \times 2 = 1441200$ bps = 1441.2 kbps. Oznacza to, że zwykła płyta CD 650MB jest w stanie zapisać 74 minuty takiego nagrania, występują jednak wersje pozwalające zapisać na niej nawet 99 minut (przy przekroczeniu standardów).

Super Audio Compact Disc (SACD)

Standard SACD został zaprojektowany z myślą zwiększenia jakości oferowanej przez nośniki typu CD-Audio, przy zachowaniu z nimi kompatybilności. Wersja finalna cechuje się dwiema warstwami, gdzie jedna jest standardowym nośnikiem CD-A, a na niej jest naniesiony przezroczysty materiał zapisujący z jakością HD dane muzyczne, które potrafią odczytać specjalnie przystosowane czytniki płyt.

Warstwa kompatybilna z CD-DA wykazuje takie same cechy techniczne jak CD-DA (kodowanie PCM, dźwięk 2.0). Warstwa HD obsługuje tylko i wyłącznie format kodowania DSD. W przeciwieństwie do PCM, cechuje się on zmienną częstotliwością próbkowania (do prawie 3 MHz), co pozwala na zdecydowanie lepsze odwzorowanie analogowego sygnału. Obsługuje ona też zapis wielokanałowy, do 6 dyskretnych kanałów audio (w systemie 5.1).

Jest to więc nieporównywalnie lepszy nośnik do zapisu audio od CD-DA, jednak nasycenie rynku jest niskie. Spowodowane jest to małą ilością wydanych płyt, a także (a może przede wszystkim) niskim poziomem świadomości przeciętnego użytkownika.

DVD-Audio (DVD-A)

Konkurencyjny do SACD, format ten zapisuje dane na płytach DVD. Cechuje się zależną od zapisu częstotliwością próbkowania (od 44,1 do 192kHz) oraz rozdzielczością (16, 20 lub 24 bity). Podobnie jak format SACD, potrafi oferować znacznie wyższą jakość względem płyt CD-Audio, przy czym potrafi zapisywać ze zmienną ilością kanałów od 1.0 do 5.1. Należy zauważyć, że na jednym dysku DVD-A może znajdować się wiele ścieżek zapisanych z różną jakością.

Format plików audio to PCM, który jest niekiedy kompresowany za pomocą bezstratnego algorytmu MLP (nazwa sumaryczna PPCM).

Ten standard nie jest kompatybilny z CD-DA, cechuje się jednak nieco wyższą fleksyjnością względem SACD. Obecnie na rynku panuje wojna pomiędzy obydwojema formatami, która nie jest jednak tak nagłaśniana jak starcie blu-ray i hddvd. Jest to częściowo spowodowane nie tylko niewiedzą, ale też sceptycyzmem niektórych użytkowników odnośnie odczuwalnie wyższej

jakości dźwięku w takim typie zapisu danych.

2.1.2 Standardowe formaty muzyczne

Chociaż współczesne komputery mogą bez problemu zapisać na swoich dyskach twardych tysiące nieprzetworzonych płyt CD-Audio, to chcielibyśmy zachować na nich jeszcze większą ilość muzyki. Należy więc wprowadzić pewnego rodzaju kompresję, która pozwoli na lepsze wykorzystanie pamięci, a także (w przypadku przesyłania plików przez sieć) szybszą ich wymianę. W dzisiejszych plikach audio najczęściej stosuje się kompresję stratną. Do stratnych formatów zapisu dźwięku należą: MP3, MP3PRO, WMA, Vorbis, AC3, VQF, RealAudio, DTS czy QTA.

Jakość spakowanej muzyki w wyżej wymienionych formatach jest dla zwykłego człowieka korzystającego ze standardowych głośników podobna do tej z płyt CD. Sprawdza się to szczególnie w muzyce pop oraz rock, a także techno, jednak zauważalne niedociągnięcia pojawiają się przy muzyce takiej jak jazz czy muzyka symfoniczna.

MPEG-1 Audio Layer-3 (MP3)

Format MP3 jest obecnie najczęściej spotykanym typem plików audio. Wykorzystuje on zmodyfikowaną dyskretną transformatę cosinusową oraz model psychoakustyczny do kompresji i kodowania muzyki. Może też wykorzystywać metodę zmiennej przepływności (VBR), dzięki której więcej miejsca przeznaczają się na bardziej skomplikowane fragmenty. Pozwala to na uzyskanie większej jakości pliku przy analogicznej jego wielkości. Sposób kompresji zdecydowanie przypomina format JPEG.

Omawiany format został opublikowany w roku 1994. Popularność zyskał jednak dopiero kilka lat później, kiedy komputery domowe bez problemu radziły sobie z jednoczesnym dekodowaniem muzyki w tym formacie i pracą biurową w czasie rzeczywistym.

Format MP3 cechuje się zmiennym bitrate, które może wynieść w zależności od kompresji od 8kbps do 320 kbps. Standardowy format MP3 pozwala na zapis stereo; wielokanałowy jest możliwy przy użyciu rozszerzenia (stworzonego przez twórców formatu) o nazwie MP3 Surround.

MP3PRO

MP3PRO jest to kodek, który powstał na bazie MP3. Pozwala on na zmniejszenie objętości plików dźwiękowych o około 50%, przy jednoczesnym nieznacznym zwiększeniu jakości.

Wykorzystuje on standardowe metody kompresji MP3, wsparte algorytmem SBR, który pozwala na zapisywanie plików z mniejszym wskaźnikiem przepływności względem MP3, jednak z porównywalną (niekiedy wyższą) jakością.

Format MP3PRO może być odtworzony za pomocą dowolnego czytnika

MP3, jednak, aby wykorzystać jego wszystkie możliwości, potrzebny jest specjalny plugin lub odtwarzacz.

Windows Media Audio (WMA)

Format kompresji dźwięku WMA został stworzony przez Microsoft, a jego najnowsza wersja 9 pojawiła się w roku 2003. Tak na prawdę standard WMA składa się z czterech różnych kodeków: WMA, WMA Pro, WMA Lossless oraz WMA Voice.

WMA jest najczęściej używanym kodekiem z tej czwórki. Jej sample cechują się podwyższoną częstotliwością (48kHz) w trybie stereo. W najnowszej wersji obsługuje tryby VBR oraz ABR. Kodowanie bazuje na algorytmie zmodyfikowanej dyskretnej transformaty cosinusowej (MDCT), który dzieli strumień bitowy na ramki i transformuje je do domeny częstotliwości, gdzie następuje kwantyzacja, po czym wynik jest poddany kompresji Huffmana.

WMA Professional jest nieco bardziej zaawansowanym kodekiem. Bazuje on na nowym algorytmie kompresji, który pozwala na zwiększenie jakości także przy niskim bitrate. Pozwala także na zapis audio z rozdzielczością 24bit, 92kHz w ośmiu kanałach (7.1). Microsoft twierdzi, że sam format technicznie jest w stanie obsłużyć nielimitowaną ilość kanałów. WMA Pro z racji bycia młodszym bratem WMA, posiada nadal małe wsparcie ze strony oprogramowania.

WMA Lossless wykorzystuje VBR i pozwala na bezstratny zapis muzyki. Został stworzony głównie z myślą o archiwizowaniu wykonanych ścieżek dźwiękowych. Jego charakterystyka to maksymalnie 96 kHz, 24bit i 6 kanałów (5.1).

WMA Voice został stworzony z myślą o przesyłaniu głosu człowieka. Jego nominalna częstotliwość to 22.05kHz, mono. Wykorzystuje CBR z bitrate dochodzącym do 20kbps. Głównym powodem stworzenia tego formatu jest fakt, że standardowy WMA źle radzi sobie z zapisem głosu.

Dolby Digital Audio Compression-3 (AC3)

Dolby Digital jest to stratny kodek wykorzystujący zapis PCM z częstotliwością 48kHz i próbkowaniu 16 lub 24bit, obsługuje dźwięk wielokanałowy. Podobnie jak WMA występuje w wielu wersjach, a każda z nich jest przystosowana do konkretnego zastosowania (np. Dolby Digital Live dla technologii interaktywnych, czy Dolby TrueHD dla dźwięku w płytach HD DVD). Pojedyncza ramka dźwięku ma wielkość 1536, składając się z sześciu sampli po 256.

Advanced Audio Coding (AAC)

AAC to stratny format zapisu plików audio. Znany jest przede wszystkim z tego, że jest wykorzystywany przez iTunes oraz urządzenia takie jak iPod czy Creative Zen. Obsługuje on próbkowanie od 8 do 320 kbps, i potrafi zapisać do sześciu kanałów.

Format ten jest uważany za następcę MP3, oferuje bowiem zapis podobnej jakości dźwięku w mniejszym rozmiarowo pliku.

Digital Theatre Systems (DTS)

DTS jest systemem kompresji dźwięku, który oferuje wyższą jakość dźwięku względem AC3 dzięki wyższemu bitrate (do 1.5MBps) oraz mniejszemu stopniowi kompresji.

Obsługuje on do ośmiu kanałów audio i częstotliwość 192kHz. Wykorzystywany jest głównie jako element ścieżek dźwiękowych, np. w wyższej jakości filmach DVD, czy też w kontenerze MKV.

Dzięki uproszczeniu dekodowania dźwięku względem AC3, zużywa on mniej zasobów i jest znacznie bardziej użyteczny w takich zastosowaniach jak gry video.

RealAudio (RA)

Przedstawiony w roku 1995 przez firmę RealNetworks (najnowsza wersja ma numer 10 i pochodzi z roku 2006), RealAudio został stworzony z myślą o wykorzystaniu w Internecie. Do jego odbioru (w niskiej jakości) wystarczy modem o prędkości 28 800 bodów - wykorzystując strumieniowanie dźwięku. Użyte kodeki zależą od wersji. W najnowszej, wykorzystywany jest autorski system RealAudio Lossless. Pozwala on na nadawanie na żywo, osiągając jakość radiową już dla gęstości rzędu 30 kbps.

Charakteryzuje się on też dużą szybkością kodowania i dekodowania, nie wymaga więc dużych zasobów systemowych do jego zapisywania i odtwarzania w czasie rzeczywistym.

Pliki RealAudio cechują się mniejszym rozmiarem niż analogiczne pliki MP3 i są gotowe do strumieniowania bezpośrednio po ich zakodowaniu (i stworzeniu pliku metadanych).

2.1.3 Kontenery multimedialne

Kontener multimedialny jest formatem pliku, który zawiera różne typy danych, takie jak wiele strumieni obrazów, wideo, audio, napisy itp. Najbardziej znane kontenery audio to WAV, AIFF i XMF. Inne (zwykle AV) to m.in. 3GP, ASF, AVI, Matroska, OGG i RealMedia. Poniżej przedstawię niektóre z nich.

WAVEform audio format (WAV)

Format WAV został stworzony przez IBM i Microsoft; jest to bezstratny

format zapisu plików audio. Choć obsługuje dowolne kodeki audio, to zazwyczaj stosuje się format PCM, co sprawia, że pliki zajmują dużo miejsca. Najwyższa jakość obsługiwana przez ten format to 44.1kHz, 16bit, stereo, czyli ma takie same parametry jak płyta CD-DA. W celu zmniejszenia rozmiaru pliku można zmniejszyć jakość, nawet do 8kHz, 8bit, mono - taki plik wynikowy zajmował będzie prawie 22 razy mniej.

Do nagrania dźwięku w takim formacie, można użyć Rejestratora dźwięków, który znajduje się w każdym systemie Windows. Od wersji 2000 został wprowadzony nowy znacznik, pozwalający na wielokanałowe nagrywanie dźwięku. Sam format jest popularny w systemach Windows oraz w przenośnych urządzeniach audio (nieskomplikowane konwertowanie do takiej postaci). Bardzo podobny do formatu WAV jest Audio Interchange File Format (AIFF), stosowany w komputerach Macintosh, oraz Interchange File Format (Amiga).

3GP

3GP jest to kontener multimedialny wykorzystywany w telefonii 3G i opiera się na formacie MP4. Przechowuje strumienie wideo i audio w sposób taki, który pozwala na szybkie dekodowanie ich przez urządzenia charakteryzujące się słabą mocą, takie jak telefony komórkowe.

Matroska (MKV)

Matroska jest darmowym kontenerem, który powstaje w koncepcji open standard i open source. Charakteryzuje się tym, że jest bardzo elastyczna, i nie ma z góry założonych ograniczeń dla poszczególnych komponentów audio czy video. Najczęściej jednak posiada ona w sobie film zakodowany w formacie H.264 (dawniej RealVideo10), muzykę w AC3 lub DTS oraz napisy w formacie ASS lub SSA.

W szczególności, ten kontener jest w stanie przetransmitować dla jednego filmu kilka różnych ścieżek dźwiękowych i wizualnych, będących komplementarnymi względem siebie. Także, podczas tworzenia pliku w formacie mkv, można zdefiniować opóźnienie ścieżki dźwiękowej.

RealMedia (RM)

RealMedia to format kontenera multimedialnego opracowanego przez RealNetworks, zwykle służy do przesyłania koherentnych strumieni składających się z RealAudio oraz RealVideo.

Ponieważ pozwala on przechowywać tylko strumienie danych typu CBR, istnieje rozszerzenie tego kontenera, które wykorzystuje VBR - RealMedia Variable Bit Rate (RMVB).

Kodeki RM (jak i RMVB) do odtworzenia wymagają specjalnych dekodek lub kompatybilnych aplikacji. Te zwykle są stworzone stosunkowo nieefektywnie, a sama jakość źródeł nierzadko pozostawia wiele do życzenia, co

sprawia, że wykorzystanie tego formatu ogranicza się w zasadzie tylko do zastosowań low-end.

2.1.4 Pliki typu Musical Instrument Digital Interface (MIDI)

Osobną kategorią jest typ MIDI. Początkowo został on stworzony jako interfejs do przesyłania informacji między elektronicznymi instrumentami muzycznymi, korzystając z pięciopinowego złącza typu MIDI.

Ogólnie, pozwala on na odgrywanie 87 różnych dźwięków jednocześnie dla szesnastu kanałów. Zakres ten można przełożyć na dźwięki pochodzące z instrumentów klawiszowych w odniesieniu do c^1 o pięć oktaw poniżej i powyżej, wliczając w to półtony. Konsekwencją tego jest fakt, że na różnych syntezatorach plik MIDI może być odtwarzany w różny sposób, a twórca zna tylko przybliżony jej kształt.

Pomimo możliwości odegrania tylko prostego dźwięku przez ten format, pozostaje on w szerokim użyciu. Na przykład, mając na względzie wielkość takiego pliku (kilka minut muzyki potrafi zająć 30 kB) na jego podstawie dodaje się prosty podkład muzyczny na strony WWW. Wykorzystuje się go także do plików Karaoke, dzięki jego prostocie i małej wadze.

2.2 Aplikacje do kodowania plików audio

Na rynku dostępne jest wiele programów, które pozwalają na przekodowywanie muzyki z jednego formatu na inny. Najwięcej z nich jednak pracuje na formatach typu MP3 - choćby z tego powodu, że są one najpowszechniejsze. Jedną z nich jest *dBpowerAMP*. Po uiszczeniu odpowiedniej opłaty, pozwala on na podstawową pracę z samym plikiem, jak na przykład wyciszanie końcówek, praca na kanałach, tworzenie karaoke itp. Jednak w wersji freeware otrzymujemy bardzo dobrą aplikację do zmiany formatu plików audio, która obsługuje dużą ilość kodeków.

Podobnie *Cool Audio Video Converter*, który jest bardzo prostym programem służącym do konwertowania i wydobywania plików audio z kontenerów video i zmiany ich formatu.

Z darmowych aplikacji możemy skorzystać z *Audacity* (który też pozwala nam nieco obrobić obsługiwany plik) lub *GX::Transcoder*, który potrafi obsłużyć blisko czterdzieści formatów audio z odpowiednimi parametrami.

W kolejnej części będę wykorzystywał jednak program *Easy CD-DA Extractor 11.53.1*, z racji prostej obsługi i wysokiej kultury generowania plików wyjściowych.

3 Ilustracja cech wybranych metod kodowania

Testy zostały wykonane na komputerze z procesorem AMD Athlon 64x2 6000+, kartą dźwiękową Sound Blaster Live! 5.1 oraz głośnikach Tracer 2.1. Do rpiu i konwersji został we wszystkich przypadkach użyty program *Easy CD-DA Extractor 11.53.1*.

3.1 Kodowanie muzyki pop

W tym przypadku źródłem będzie fragment z płyty CD-DA STING "Brand New Day" o nazwie "Desert Rose" (czas trwania: 4:46).

Wykonywanie kopii muzyki z CD-DA jest jednym z najczęstszych zastosowań praktycznych, które są poruszane w tym laboratorium.

Przebieg konwersji oraz wyniki przedstawia poniższa tabelka.

Format	Bitrate	Częstotl.	Kanały	Rozmiar [b]
MP3	320 kbps	44,1 kHz	2	11 446 367
MP3 (Near-CD)	192 kbps	44,1 kHz	2	6 869 714
MP3	128 kbps	44,1 kHz	2	4 581 388
MP3 VBR-5	112-192 kbps	44,1 kHz	2	5 130 215
MP3 ABR-160	128-160 kbps	44,1 kHz	2	5 601 181
WMA Lossless	969 kbps	44,1 kHz	2	34 621 026
WMA	320 kbps	44,1 kHz	2	11 487 140
WMA	128 kbps	44,1 kHz	2	4 606 928
WMA (Near-CD)	64 kbps	44,1 kHz	2	2 321 492
WMA VBR-5	80 kbps	44,1 kHz	2	2 920 566
WMA VBR-9	183 kbps	44,1 kHz	2	6 556 622
AAC	320 kbps	44,1 kHz	2	11 439 326
AAC	128 kbps	44,1 kHz	2	4 575 909
AAC (Near-CD)	112 kbps	44,1 kHz	2	4 004 160
Format	Rozdz.	Częstotl.	Kanały	Rozmiar [b]
WAV (CD)	16bit	44,1 kHz	2	50 443 666
WAV	16bit	44,1 kHz	1	25 221 994
WAV	8bit	44,1 kHz	2	22 876 550
WAV	8bit	44,1 kHz	1	12 611 158

Czas wykonywania wszystkich kompresji wynosił kilka sekund (z wyjątkami takimi jak AAC w trybie 320kbps).

Z niewiadomych przyczyn pliki w formacie AAC miały długość 4:50 względem pozostałych 4:45.

We wszystkich testowanych formatach próbki Near-CD oraz lepsze podczas odtwarzania dźwięku cechowały się niskim poziomem różnicy jakości względem płyty CD. Jednak podczas odsłuchiwania próbek 320kbps oraz 128kbps

zakodowanych w MP3, WMA oraz AAC można było wyczuć subtelną różnicę w dźwięku.

Z tego powodu na wybór odpowiedniego kodeka największy wpływ powinien mieć rozmiar pliku wynikowego w żądanej jakości. Z przetestowanych w tej podsekcji formatów, spośród trybów Near-CD, najlepiej poradził sobie WMA, zajmując 2 321 492 bajty.

W porównaniu do AAC oraz WMA, MP3 poradziło sobie z kompresją słabo. Wprawdzie wielkość pliku wynikowego dla bitrate 128kbps była prawie taka sama dla wszystkich trzech, jednak porównywalnie jakość zaoferowana przez MP3 okazała się najniższa.

Przedstawiony tutaj raczej tylko dla porównania format WAV w jakości CD cechował się podobną do innych formatów jakością, a jego okrojone wersje (pojedynczy kanał, mniejsza rozdzielczość) cechowały się zdecydowanym spadkiem jakości. Wygenerowane pliki pozwalają jednoznacznie stwierdzić, że format ten w zasadzie nie ma już dziś żadnego zastosowania w tym typie dźwięku - nawet bezstratnie zapisywany WMA cechuje się mniejszą wielkością niż WAV. Jest on wprawdzie szybciej odczytywany, jednak dostępne dziś zasoby sprawiają, że ten fakt jest pomijalny.

3.2 Kodowanie muzyki klasycznej

Wolfgang Amadeusz Mozart - Eine kleine Nachtmusik - ten utwór został przeze mnie wykorzystany do przetestowania cech przekodowywania plików audio dla muzyki klasycznej. W tym przypadku, źródłem był plik mp3, o czasie trwania 4:16 oraz stosunkowo wysokiej jakości - 256kbps, 44kHz, stereo.

Ponieważ plik wyjściowy był w jakości 256kbps, otrzymane pliki w wyższej jakości nie mogły wykazywać się wyższą kulturą odbioru niż plik wyjściowy, więc zrezygowałem w tym punkcie z generowania takich próbek.

Konwersja została wykonana dla analogicznych ustawień jak w sekcji 3.1 dla muzyki pop. Wielkości plików wynikowych okazały się w zdecydowanej większości przypadków ok. 8-15% mniejsze, co jest proporcjonalnym spadkiem w porównaniu do czasu trwania utworu.

Ponownie format WMA okazał się stosunkowo najlepszy w relacji jakość - wielkość. Względem wyjściowych 8 209 723 bajtów wielkości pliku MP3, plik WMA o akceptowalnym w tym przypadku wydźwięku zajmował jedyne 1 570 054 bajtów.

Na tym tle, WMA rysuje się nam jako najlepszy kodek do zastosowań domowych z wybranych na początku pracy formatów. Przejdźmy do kolejnego kroku, czyli kodowania głosu ludzkiego.

3.3 Kodowanie mowy

We współczesnym internecie coraz częstszą metodą komunikacji jest głos. O ile jednak wpływ na wybór kodeka przy wykorzystywaniu takich aplikacji jak skype czy googletalk mamy żaden, a w bardziej zaawansowanych programach typu ventrilo - niewielki, to coraz częstszą formą tworzenia takiego contentu jak blogi są podcasty, czyli treść notki przekazana w formie nie pisanej, ale mówionej (przez pliki video lub audio). Zagadnienie sprowadza się do oczywistego pytania o najlepszy kodek, który przy odpowiednich parametrach zapewni nam odpowiednio wysoką jakość przy utrzymaniu małego rozmiaru pliku.

Należy tutaj zwrócić uwagę na to, że akceptujemy niższą jakość pliku w tym przypadku, pod warunkiem, że jesteśmy w stanie bez problemu zrozumieć tekst, oraz nie jest on tak "płytki", że słuchanie go sprawia ból dla uszu.

W celu przetestowania jakości kodowania mowy przez różne kodeki został nagrany, używając rejestratora dźwięku, krótki, dwusekundowy plik audio, zapisany w formacie WMA - taki plik w jakości 96kbps/44kHz/stereo wykorzystał 40891 bajtów pamięci. Przebieg konwersji oraz wyniki przedstawia poniższa tabelka.

Format	Bitrate	Częstotl.	Kanały	Rozmiar [b]
MP3	128 kbps	44,1 kHz	2	38 161
MP3	40 kbps	16 kHz	2	11 828
WMA	20 kbps	22 kHz	1	13 920
WMA-Talk	20 kbps	22 kHz	1	12 020
AAC	256 kbps	44 kHz	2	73 689
AAC	40 kbps	32 kHz	2	11 852
AAC	20 kbps	44 kHz	1	6 117

Z przedstawionych konwersji, akceptowalne do odsłuchu były następujące: WMA 20kbps, WMA-Talk 20kbps, MP3 128kbps, AAC 40kbps, AAC 256kbps. Po odrzuceniu nadmiarowych formatów (z bitrate większym od 100kbps) pozostały WMA, WMA-Talk oraz AAC.

W przypadku formatu WMA-Talk dźwięk był nieco zmodyfikowany względem oryginału, jednak można było go odsłuchać bez problemu, ze zrozumieniem tekstu. Różnica w wielkości między WMA-Talk oraz AAC była pomijalnie mała (w okolicach połowy procenta), i właśnie te formaty w przedstawionej jakości powinny być używane do kodowania i przesyłania tekstu mówionego.

4 Podsumowanie

W pracy zostały przedstawione formaty zapisu i kodowania plików audio. Niektóre z nich (te bardziej znane) zostały przetestowane. Wyniki testów

praktycznych pokazały, że przyzwyczajenie użytkowników do formatu MP3 jest w chwili obecnej irracjonalne, gdyż analogiczną jakość można uzyskać zapisując w innym formacie znacznie mniejszą wielkością plik. Przede wszystkim mam tutaj na myśli format WMA. Wydaje się, że w zasadzie jedyną przeszkodą w rozwoju tego formatu i zwiększenia jego udziału w rynku są przenośne odtwarzacze muzyki, które zwykle obsługują tylko format MP3. Chociaż w artykule prawie nie został poruszony temat zapisu audiofilskiego muzyki, to na rynku w zasadzie nie istnieje format kompresji bezstratnej dźwięku (może poza formatem FLAC). Znaczna większość algorytmów kompresujących jest nastawiona na upakowanie danych w sposób stratny, co powoduje utratę jakości audio. I w zasadzie żaden z formatów nie reklamuje się jako "nowocześnie audiofilski", może poza wspomnianymi formatami zapisywania dźwięku na płytach SACD/DVDA. W zamian, prawie wszystkie kodowania sprowadzają się do kompresji muzyki 44,1 kHz / 16 bit. W ciągu ostatnich lat bardzo duży rozwój w zakresie plików audio dokonał się także na polu formatów kontenerów multimedialnych. Chyba najnowocześniejszym takim formatem jest MKV, który umożliwia nawet tworzenie ripów muzyki i filmu z płyt Blu-Ray.

Literatura

- [1] <http://www.wikipedia.org/>
- [2] <http://www.fizykon.org/>
- [3] <http://www.cdrinfo.pl/>
- [4] <http://www.pwii.pl/>
- [5] <http://www.germanixsoft.de/>
- [6] <http://sudacity.sourceforge.net/>
- [7] <http://www.dbpoweramp.com/>
- [8] <http://bg.pwr.wroc.pl/>
- [9] <http://dariuszpod.republika.pl/>
- [10] <http://www.pcworld.pl/>