

Sprzętowy dekodery MP3, część 1

Muzyka zapisana w plikach MP3 podbija świat. Pomimo wielu zastrzeżeń związanych z jakością odtwarzanej muzyki popularność formatu MP3 rośnie i nic nie wskazuje, że ekspansja zostanie zatrzymana. Biorąc pod uwagę ten fakt postanowiliśmy przybliżyć naszym Czytelnikom konstrukcję sprzętowego dekodera MP3 wykonanego na bazie jednego z łatwiej (co nie oznacza, że łatwo!) dostępnych dekoderek - STA013 firmy STMicroelectronics.

W odróżnieniu od większości przygotowywanych przez nas projektów, urządzenie prezentowane w artykule nie znajdzie się w ofercie kitów. Jest to wynik wielu problemów zaopatrzeniowych, na jakie natknęliśmy się podczas przygotowywania prototypów.

Naszym pierwotnym zamierzeniem było opracowanie przenośnego odtwarzacza MP3 zintegrowanego z czytnikiem kart pamięciowych Flash w standardzie MMC (ang. MultiMedia Card), co w znacznym stopniu się powiodło, ale nie zakończyło 100-procentowym sukcesem. Niestety, większość podzespołów niezbędnych do wykonanie dekodera jest do-

stępna tylko w ilościach przemysłowych, co uniemożliwiło - przynajmniej na razie - wdrożenie kitu do produkcji.

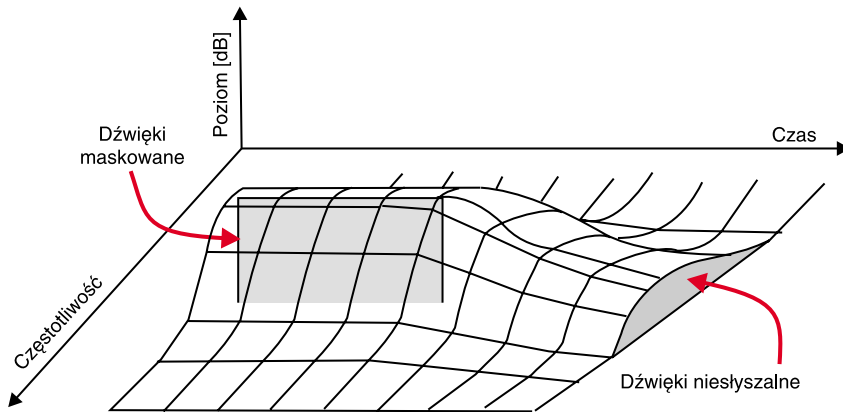
Pomimo tych problemów postanowiliśmy przedstawić opis tej wersji dekodera, którego konstrukcja została nieco zmodyfikowana, dzięki czemu urządzenie można wykorzystać także jako sprzętowy dekodery MP3 współpracujący z komputerem PC lub - po adaptacji oprogramowania - z dowolnym innym. Prowadzimy dalsze prace mające na celu zastąpienie pamięci MMC dyskiem twardym z interfejsem ATA. O efektach będziemy informować.

MP3 - co to jest?

Mianem MP3 określane są pliki zawierające cyfrowo zapisaną muzykę, których najpoważniejszą zaletą są niewielkie objętości, co najmniej kilkukrotnie mniejsze niż wynika to z zastosowanych metod konwersji A/C. Algorytm zastosowany do kodowania i dekodowania sygnału zaczerpnięto ze standardu MPEG (jego wersje 1/2/2,5), gdzie zajmuje on w „stosie“ kodeka warstwę 3, odpowiadającą za obróbkę sygnałów audio. Założeniem przyświecającym twórcom tego standardu było zapewnienie jakości odtwarzanej muzyki porównywalnej z CD przy jednoczesnym obniżeniu przepływności danych w kanale cyfrowym.

W zależności od oczekiwanej jakości odtwarzanego sygnału w MP3 można zmieniać częstotliwość próbkowania, liczbę kodowanych kanałów (od 1 do 9) oraz liczbę bitów odpowiadających pojedynczej próbce sygnału audio. Dobór tych parametrów ma oczywiście ogromny wpływ na wymaganą przepływność toru cyfrowego, ale sukces nadzwyczajnej kompresji sygnału audio (od 2,7:1 aż do 24:1) jest ukryty gdzie indziej





Rys. 1. Ilustracja zasady maskowania dźwięków.

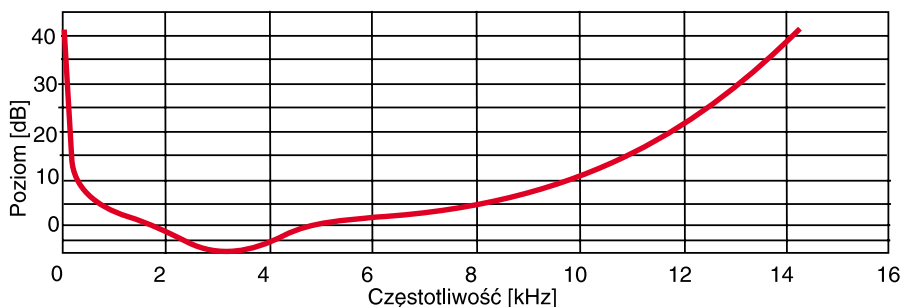
- w sprytnym wykorzystaniu niedoskonałości ludzkiego słuchu. Oznacza to rezygnację z dążenia do idealnego odtworzenia sygnału audio i zastąpienie tego dążenia wytworzeniem przez dekoder sygnału wyglądającego z punktu widzenia ludzkiego ucha tak, jak sygnał oryginalny. Zgodnie z informacjami propagowanymi przez kompetentne instytucje zastosowanie współczynnika kompresji 6:1, przy 16-bitowej próbce i częstotliwości próbkowania 48kHz, nie powoduje zmiany w jakości, możliwej do wychwycenia przez ekspertów odsłuchowych. Warto zwrócić uwagę, że jakość tak zakodowanego sygnału jest nieco lepsza niż w przypadku CD, a konieczna przepływność danych wynosi zaledwie 256kb/s zamiast 1,4Mb/s (w przypadku CD bez kompresji).

Uzyskanie tak znacznego ograniczenia wymaganej przepływności jest możliwe po zastosowaniu kilku „chwytów“. Analogowy sygnał wejściowy jest analizowany przez koder MP3 pod kątem możliwości pominięcia niektórych fragmentów sygnału audio. W procesie maskowania wykorzystano

właściwość ludzkiego słuchu polegającą na maskowaniu (braku słyszenia) sygnałów o mniejszych amplitudach przez sygnały o amplitudach znacznie większych (rys. 1). W celu zwiększenia precyzji działania algorytmu kompresującego, całe pasmo audio podzielono na 32 podpasma, z których każde obrabiane jest w oddzielnym torze audio.

Drugą istotną właściwością ludzkiego słuchu, wykorzystaną przez twórców standardu MPEG, jest próg słyszalności i czułości na zniekształcenia, zależny od średniego natężenia otaczających dźwięków i ich widma. Dzięki temu, dźwięki „ukryte“ pod krzywą pokazaną na rys. 2 nie muszą być kodowane, ponieważ i tak nie będą słyszane. W przypadku kompresowania rozbudowanych pasażów muzycznych ich jakość mogłaby ulec obniżeniu, czemu zapobiega mechanizm modyfikowania progów maskowania, często nazywany rezerwuarem bitów.

Niebagatelne zmniejszenie wymaganej przepływności sygnału audio uzyskano także dzięki specjalnemu kodowaniu sygnałów stereofonicznych, który określono



Rys. 2. Krzywa czułości podprogowej ludzkiego słuchu.

mianem *Joint Stereo*. W jego ramach zintegrowano dwa mechanizmy:

- Intensity Stereo - umożliwiający kodowanie niektórych sygnałów (szczególnie o najniższych i najwyższych częstotliwościach) w jednym kanale monofonicznym, który jest „rozplątany“ do postaci stereo przez dekoder,
- Mid/Side Stereo - pozwalający obniżyć liczbę niezbędnych bitów w przypadku znacznego podobieństwa informacji przekazywanych w kanale lewym i prawym.

Ostatnim etapem kompresowania sygnału audio jest kodowanie zgodnie z algorytmem Huffmana, które polega na zastąpieniu długich i często powtarzających się sekwencji bitów skróconymi znacznikami, których odpowiedniki są odtwarzane w dekodерze. Czytelników zainteresowanych szczegółami związanymi z kodowaniem Huffmana zachęcamy do odwiedzenia stron w Internecie pod adresami: <http://www.compressconsult.com/huffman/#huffman> oraz <http://www.cs.duke.edu/csed/poop/huff/info/>.

Dekodowanie i dekompresja sygnału audio MPEG-L3 jest możliwe na drodze programowej lub sprzętowej. Dekodowanie progra-

Tab. 1. Sposób zakodowania numeru wersji MPEG.

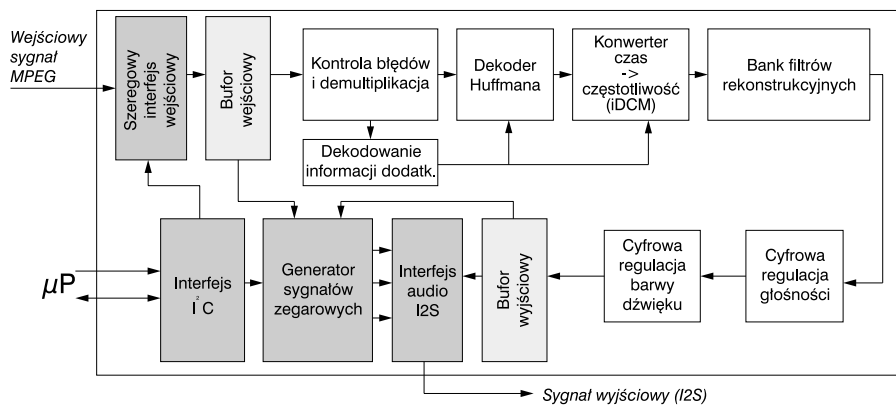
bit20	bit19	Wersja MPEG
0	0	MPEG 2.5
0	1	zarezerwowane
1	0	MPEG 2 (ISO/IEC 13818-3)
1	1	MPEG 1 (ISO/IEC 11172-3)

Tab. 2. Sposób zakodowania informacji o rodzaju sygnału.

bit7	bit6	Rodzaj sygnału
0	0	Stereo
0	1	Joint stereo (Stereo)
1	0	Dual channel (2 kanały mono)
1	1	Single channel (Mono)

Tab. 3. Sposób zakodowania informacji o emfazie odtwarzanego sygnału.

bit1	bit0	Emfaza
0	0	none
0	1	50/15 ms
1	0	zarezerwowany
1	1	CCIT J.17



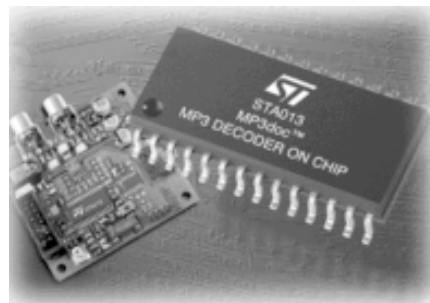
Rys. 3. Schemat blokowy sprzętowego dekodera MP3.

mowe w czasie rzeczywistym jest możliwe ze względu na stosunkowo niewielką wymaganą przepływność danych i dużą wydajność współczesnych procesorów stosowanych w komputerach (stacjonarnych i przenośnych). Czysto programowa metoda dekodowania sygnału MPEG w sprzęcie przenośnym nie jest uzasadniona, ponieważ pobór energii przez procesory uniwersalne, a także przez większość procesorów DSP jest dość duży. Lepszym wyjściem jest zastosowanie sprzętowego dekodera (np. STA013), którego schemat blokowy pokazano na rys. 3, a na rys. 4 został przedstawiony algorytm jego pracy.

Format plików MP3

Pliki MP3 mają strukturę opisaną w normach MPEG (m.in. w ISO11172-3). Plik rozpoczyna się zawsze od 4-bajтового nagłówka, w którym pierwsze 11 bitów (31...21) o wartości „1” umożliwia zsynchronizowanie odbiornika z nadajnikiem, kolejne dwa bity (20, 19) zawierają informację o wersji MPEG (tab. 1), bity 18 i 17 zawierają informację o numerze warstwy MPEG (dla L3 przyjmują one stany: 01), kolejny bit (16) niesie informację o włączonej lub wyłączonej sprzętowej weryfikacji sumy kontrolnej CRC

(opcjonalnie dodawanej w postaci 2 8-bitowych słów na końcu nagłówka), cztery kolejne bity (15...12) informują o przepływności zakodowanego sygnału, bity 11 i 10 informują o częstotliwości próbkowania. Bit o numerze 9 informuje o zintegrowaniu w ramce danych slotu wyrównującego długość ramki do długości wynikającej z zestawienia częstotliwości próbkowania z oczekiwaną przepływnością. Kolejny, 8 bit jest oznaczony jako *Private* i jest wykorzystywany tylko w celach in-



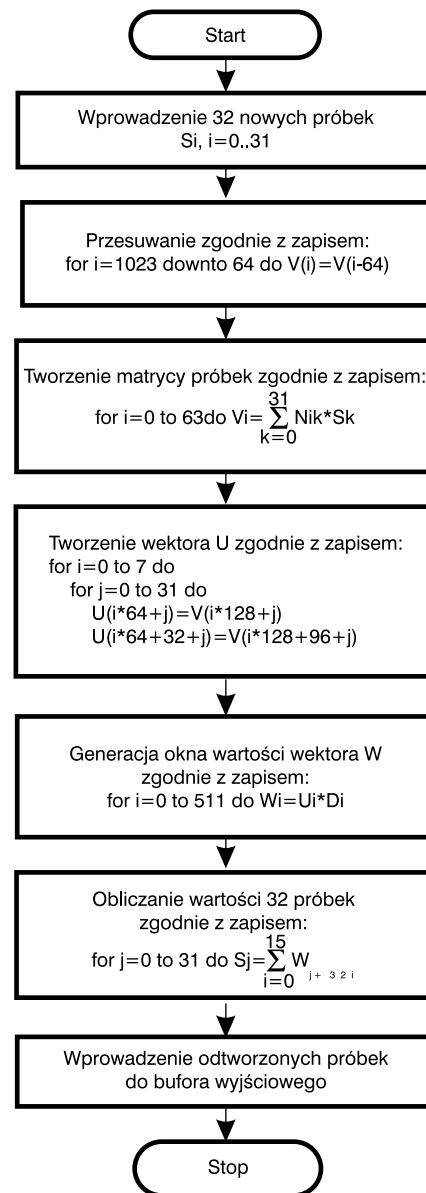
formacyjnych. Dwa kolejne bity (7 i 6) zawierają informacje o rodzaju przesyłanego sygnału (tab. 2), a dwa następne zawierają informacje niezbędne do prawidłowego dekodowania sygnałów stereofonicznych, kodowanych z wykorzystaniem metody *Joint Stereo*. Bit o numerze 3 informuje o zastrzeżeniu praw autorskich do danych zgromadzonych w pliku, natomiast bit 2 sygnalizuje czy jest to nagranie oryginalne, czy też jest to kopia oryginału. Dwa najmłodsze bity nagłówka zawierają informację o rodzaju emfazy niezbędnej do prawidłowego odtworzenia sygnału audio (tab. 3).

Po nagłówku są przesyłane dane audio, na końcu których

znajduje się 128-bitowe pole informacyjne, zawierające szczegółowe informacje o pliku audio, w tym m.in.: tytuł utworu, nazwę wykonawcy, rok publikacji itp. (tab. 4).

Pamięci MMC

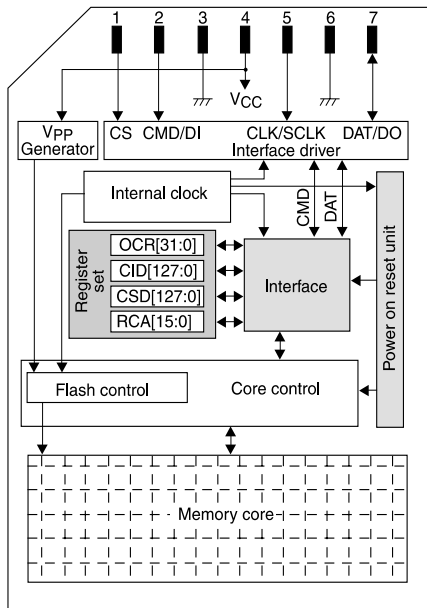
Większość przenośnych odtwarzaczy MP3 jest wyposażona w gniazdo pamięci MMC lub innych, wymiennalnych pamięci nielotnych. Większość pamięci zgodnych ze standardem MMC jest przystosowana także do wymiany danych z otoczeniem poprzez interfejs SPI, który wykorzystano w prezentowanym projekcie. Rezygnacja ze standardo-



Rys. 4. Algorytm pracy dekodera MP3.

Tab. 4. Zawartość pola informacyjnego pliku MP3 (ostatnie 128 bajtów pliku).

Tytuł utworu	30 znaków
Wykonawca	30 znaków
Nazwa albumu	30 znaków
Rok produkcji	4 znaki
Komentarz	30 znaków
Liczba utworów w pliku	3 znaki
Genre	1 znak



Rys. 5. Schemat blokowy karty MMC.

wego protokołu stosowanego w MMC wynika z faktu jego znacznego (w porównaniu z SPI) skomplikowania, a także z nieco prostszej w implementacji sprzętowej obsługi SPI.

Schemat blokowy karty MMC2-XA3 (o pojemności 16MB) firmy Infineon, aktualnie wytyczającej standard rynkowy, pokazano na rys. 5.

Jak widać na rysunku, oprócz matrycy pamięciowej Flash we wnętrzu karty zintegrowano szereg bloków niezbędnych do jej prawidłowej pracy, w tym przetwornicę napięcia programującego, programowany interfejs szeregowy, zespół rejestrów konfiguracyjnych pracę pamięci oraz zawierających informacje o parametrach karty (w tym o poborze prądu dla maksymalnego i mini-

malnego dopuszczalnego napięcia zasilającego), a także generator taktujący synchronizujący pracę wszystkich wewnętrznych mechanizmów karty. Kartę wyposażono także w blok korekcji błędów transmisji, który jest bardzo pomocny zwłaszcza podczas przesyłania danych z dużą szybkością. Dostęp do matrycy może być sekwencyjny lub swobodny, który ze względu na konieczność każdorazowego adresowania docelowej komórki pamięci nieco spowalnia zapis i odczyt.

Wprowadzenia interfejsu szeregowego są odporne na przepięcia do 4kV, a maksymalna szybkość transferu danych (odczyt) wynosi 13,7Mb/s, co z dużym zapasem wystarcza do naszych celów.

Andrzej Gawryluk, AVT