
NIECH STANIE SIĘ DŹWIĘK

PROCESOR EFEKTÓW DŹWIĘKOWYCH
DO GITARY
NA URZĄDZENIA
OPARTE NA SYSTEMIE ANDROID

Guitar effects processor for Android devices

Maciej BLASZKE, Dawid WEBER, Szymon ZAPOROWSKI

Politechnika Gdańska,
Wydział Elektroniki, Telekomunikacji i Informatyki,
maciejb93@gmail.com

STRESZCZENIE

W pracy przedstawiono procesor efektów dźwiękowych do gitary składający się z układu elektronicznego i aplikacji pracującej w czasie rzeczywistym na urządzeniach mobilnych z systemem Android. W pierwszej części przedstawiono proces przetwarzania dźwięku w aplikacji oraz interfejs użytkownika. Interfejs użytkownika napisany został w języku Java, wspartym językiem znaczników XML, zaś przetwarzanie dźwięku, ze względu na konieczność zapewnienia wysokiej wydajności, zrealizowano w języku C++. W pracy uwzględniono możliwość jednoczesnego uruchomienia kilku algorytmów połączonych w postaci bloków oraz przedstawiono sposób sterowania ich parametrami w czasie działania programu. W drugiej części przedstawiono układ (przejściówkę), dopasowujący impedancję, w postaci zasilanego baterią przedwzmacniacza, do którego podłącza się gitarę oraz urządzenie mobilne. Przeprowadzono osobno testy układu oraz aplikacji, jak i testy działania połączonego systemu układ – aplikacja.

ABSTRACT

This paper concerns the design of applications and external adapters for Android smartphones, whose task will be to implement guitar effects processor. The thesis includes a breakdown stages, from design, through implementation, analysis and conclusion ending.

The whole signal audio processing occurs in the application. It takes place on the basis of emulation of the analog effects implemented on the basis of their theoretical acting. One additional assumption is that the user can simultaneously run a number of effects and control its parameters in the real-time. The adapter is designed and constructed as a preamp powered by a 9V battery. Its main tasks are the implementation of the impedance matching between the guitar and telephone and also filtering DC component that is generated by the mobile phone as 2.2V tense phantom, which could damage the guitar's electronic components, especially with active pickups.

1. WPROWADZENIE

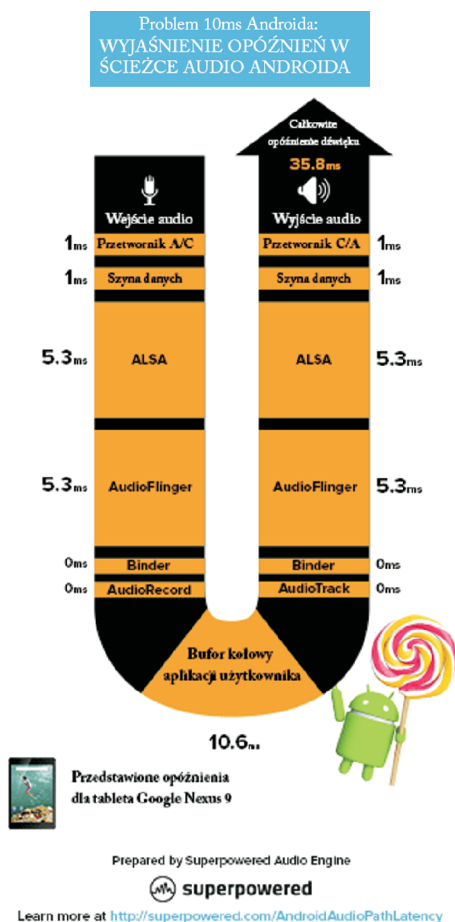
Projekt polegał na stworzeniu aplikacji procesora efektów gitarowych na platformę Android oraz specjalnej przejściówki umożliwiającej dopasowanie impedancyjne pomiędzy telefonem a gitarą. Na rynku urządzeń mobilnych można spotkać podobne rozwiązania, chociażby przeznaczone dla smartfonów firmy Samsung. Jest to jedyne tego typu urządzenie na platformę Android, cechuje się tym, że połączenie pomiędzy telefonem a gitarą korzysta ze standardu mobilnego USB. Podobnymi urządzeniami są przenośne, małe wzmacniacze gitarowe typu combo firm VOX czy też Peavey. Rozmiary tych urządzeń są na tyle małe, że można nazwać je kieszonkowymi, dzięki czemu możliwe jest granie na gitarze elektrycznej w dowolnym miejscu i o dowolnej porze.

Pomysł projektu podyktowany był brakiem uniwersalnego rozwiązania umożliwiającego zastosowanie w każdym urządzeniu z platformą Android. Istotnym aspektem była optymalizacja rozwiązań opóźnienia w torze audio platformy Android, tak aby uzyskać jak najmniejsze opóźnienia. Ze względu na ten problem postanowiono skorzystać z języka programowania C++ i natywnego środowiska NDK Crystax.

2. PODSTAWY TEORETYCZNE I PROJEKTOWE

Środowisko NDK Androida umożliwia użytkownikom przetwarzanie dźwięku, jednak jego istotną wadą są znaczne opóźnienia wprowadzane do toru audio.

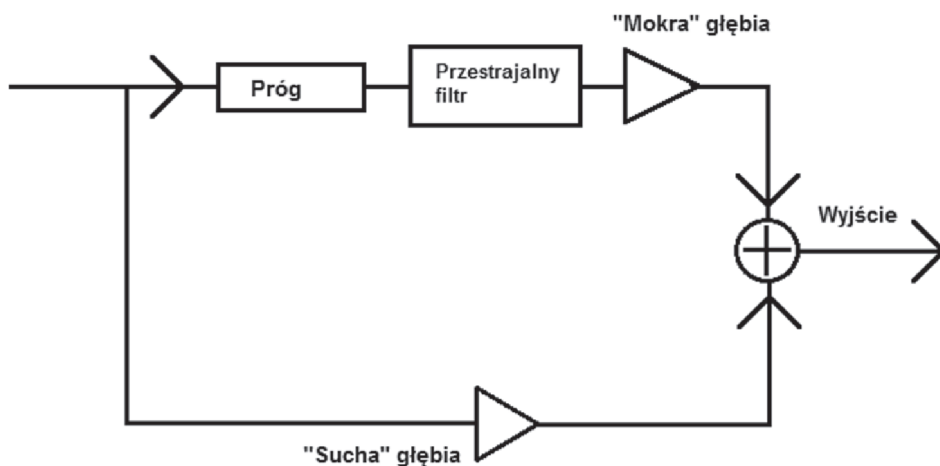
Na rysunku 1 widać, że całość toru audio jest bardzo złożona i wiele z jego elementów jest powtarzanych. Bufor HAL, czyli androidowa sprzętowa warstwa abstrakcyjna audio, „pośredniczy” pomiędzy serwerem mediów systemu Android a Linuxowym sterownikiem dźwięku. Analiza kodu szeregu implementacji HAL w wersji open source w repozytorium źródłowym Androida wykazała kilka niezrozumiałych rozwiązań wprowadzających znaczne opóźnienie oraz niepotrzebnie obciążających procesor podczas przetwarzania toru audio. Wynikały one z błędnej konfiguracji oraz nieoptymalizowanego kodu. Poprawna implementacja HAL nie powinna wprowadzać żadnych opóźnień.



Rys. 1. Schemat toru audio systemu Android [1]

Analizując cały tor audio (rys. 1), dla tabletu Google Nexus 9 uzyskano opóźnienie równe 35,8 ms, które jest akceptowane w domenie przetwarzania dźwięku [1, 2].

W celu przeprowadzenia testów zarówno aplikacji, jak i przejściówki, postarano się o cyfrowe odzwierciedlenie trzech najbardziej popularnych i podstawowych efektów gitarowych takich jak przester, echo oraz tremolo.



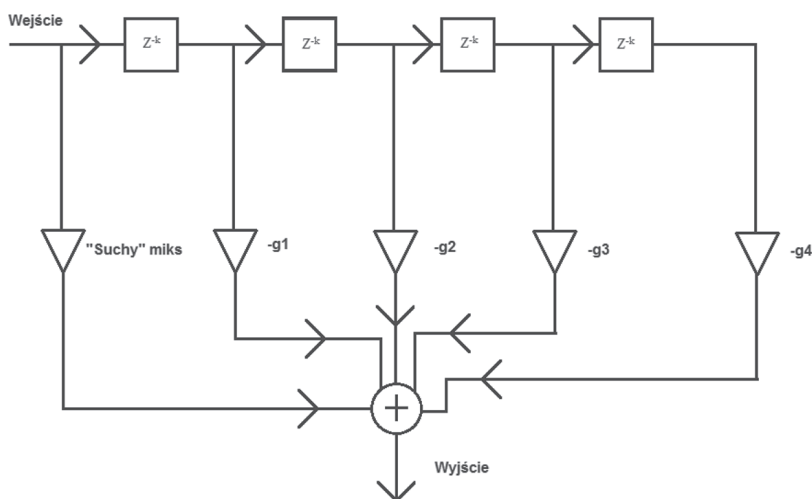
Rys. 2. Schemat blokowy efektu przesteru

Przester, którego schemat blokowy przedstawiono na rysunku 2, ma agresywne brzmienie pochodzące od wzmocnienia sygnału gitarowego, podczas którego wprowadzane są dodatkowe harmoniczne do toru oraz obcinane są wierzchołki sygnału. Cyfrowo został on zrealizowany przez wzmocnienie sygnału i ograniczenie funkcją tangensa hiperbolicznego. Poniżej podano fragment kodu w języku programowania C++ ukazujący realizację algorytmu.

```
temp_in[c] = in[c];
out[c] = (tanh((gain_*temp_in[c]) + bias_));
```

Parametry gain_, bias_ służą do zmiany właściwości efektu zgodnie z algorytmem przetwarzania. Gain_ jest wartością wzmocnienia, natomiast bias_ odpowiada za asymetryczność sygnału. Możliwości procesorów pozwalają obecnie na uzyskanie nieznacznych opóźnień, które nie są słyszalne dla człowieka. Dlatego też użycie drugiego bufora i zapisywanie do niego próbek nie powoduje dużych opóźnień. Wielkość bufora zależy od częstotliwości próbkowania i szerokości ramki [3–6].

Echo jest realizowane przez dodawanie na wejście sygnału wartości wejściowej opóźnionej o zadany czas. Można dodatkowo zrealizować je, rozbudowując o sprzężenie zwrotne, dzięki czemu można stworzyć multiecho. Na rysunku 3 przedstawiono schemat blokowy opisywanego efektu.



Rys. 3. Schemat blokowy efektu echa

Poniżej przedstawiono fragment kodu w języku programowania C++ ukazujący realizację algorytmu.

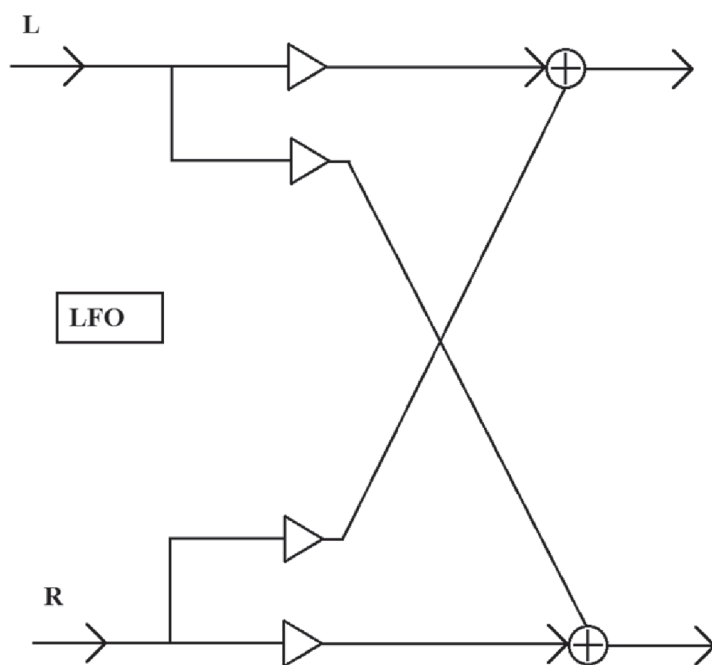
```
temp_out_[i_] = (1 - mix_)*(temp_in_[i_]) + (mix_*(temp_in_[j_++] + feedback_*temp_out_[k_++]));
out[c] = temp_out_[i_++];
if (i_ >= 100000) i_ = 0;
if (j_ >= 100000) j_ = 0;
if (k_ >= 100000) k_ = 0;
```

Efekt echa wykorzystuje bufor kołowy, gdyż algorytm wymaga użycia poprzednich próbek przy przetwarzaniu kolejnych. Sygnał wejściowy jest sumą sygnałów, w tym jednego opóźnionego. Parametr `mix_` odpowiada za obniżenie poziomu sygnału wejściowego i opóźnionego, `feedback_` jest parametrem służącym do zmiany poziomu sygnału na pętli zwrotnej [1–3, 5].

Tremolo, którego schemat blokowy przedstawiono na rysunku 4, jest efektem wpływającym na wartość amplitudy. Realizowany jest przez ograniczenie obwiedni sygnału innym sygnałem cyklicznym, na przykład sinusem, co powoduje charakterystyczne drganie dźwięku. Poniżej przedstawiono kod w języku C++ ukazujący implementację algorytmu.

```
out[c] = in[c] * (1 - depth_*0.5* sin(freq_* (i_++ * 2 * 3.14 / 44100)));
if (i_ >= 44100) i_ = 0;
```

Algorytm tremolo bazuje na okresowej zmianie amplitudy sygnału. Realizacja tej metody opiera się na zdefiniowaniu wartości zmiennych przez sinusa, który dodatkowo modulowany jest parametrem głębokości `depth_`. Głębokość ma wpływ na wzrost amplitudy przetwarzanego sygnału [1–3, 5].



Rys. 4. Schemat blokowy efektu tremolo

3. ZAŁOŻENIA PROJEKTOWE APLIKACJI

Najistotniejszą kwestią było uzyskanie jak najniższej latencji dla aplikacji. Pierwszą rzeczą, która się do tego przyczyniła, była aktualizacja Androida w postaci natywnego wsparcia obsługi czasu rzeczywistego przez wyłączenie dodatkowych buforów w HAL dla wersji systemu nowszych niż 4.4.2 (Kit Kat). Dodatkowo, aby uzyskać jak najmniejsze opóźnienie, należy wykorzystać biblioteki umożliwiające przetwarzanie audio po stronie języka C++ [7].

Język C++ jest językiem natywnym, czyli po uruchomieniu danej aplikacji napisanej w tym języku jest ona uruchomiona bezpośrednio na procesorze, co oznacza, że kompilowanie danego kodu w języku natywnym pozwala na przypisanie go konkretnemu systemowi oraz architekturze procesora.

Język Java działa na zasadzie optymalizacji kodu na daną platformę, gdzie uruchamiana jest aplikacja. Służy do tego wirtualna maszyna Javy. Mimo przenośności programów, wadą takiego rozwiązania jest dłuższe uruchamianie się. Mimo że w dłuższej działających aplikacjach, gdzie kod już został skompilowany i przeniesiony z JVM na maszynę fizyczną, różnica czasów obliczania algorytmów w języku Java i w języku C++ jest niewielka, to dla aplikacji wymagających szybszej pracy pomiędzy różnymi wątkami i przy dużych aplikacjach lepiej sprawdza się język C++. Dlatego też do napisania algorytmów przetwarzających dźwięk wykorzystano język C++ [4].

4. IMPLEMENTACJA APLIKACJI

Po stronie języka Java użyto podstawowych bibliotek służących do zbierania, przetwarzania i odtwarzania dźwięków przez mikrofon wbudowany w telefon lub przy użyciu zestawu słuchawkowego. W projekcie użyto specjalnie zaprojektowanej przejściówki, aby móc przez gniazdo typu JACK umożliwić przesłanie sygnału z przetworników gitarowych do aplikacji, następnie przetworzyć go przez odpowiednie algorytmy przetwarzania dźwięku i otrzymać już przetworzony sygnał [6].

Aplikacja opiera się na wyborze odpowiedniego pliku konfiguracyjnego, w którym zadana jest nazwa algorytmu, a następnie w kolejnych oknach użytkownik ma możliwość ustalenia wartości parametrów dla danego algorytmu.

W projekcie efektów gitarowych wykorzystano gotowe rozwiązania bibliotek przetwarzania ratlab na licencji opensource, są one napisane w języku programowania C++ w celu uzyskania szybszych metod obliczeniowych algorytmów przetwarzania sygnału audio. Klasy bibliotek posiadają informacje na temat długości bufora, liczby kanałów i częstotliwości próbkowania. Dodatkowo odpowiadają za pobieranie parametrów wybranych algorytmów. Parametry efektów podawane są bezpośrednio przez aplikację mobilną na telefonie. Po stronie części programowej w aplikacji mobilnej zdefiniowane klasy są odpowiedzialne za przekazywanie informacji o wyborze algorytmu efektu gitarowego, pobieranie i przesyłanie przez natywne metody wartości parametrów algorytmu przetwarzania.

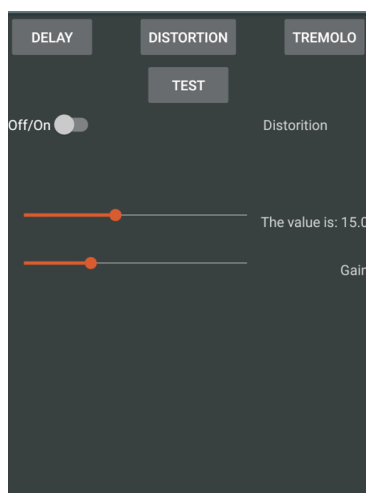
Java Native Interface (JNI) jest platformą programistyczną umożliwiającą przywołanie kodu napisanego w języku Java przez natywne aplikacje i biblioteki napisane w innych językach programowania takich jak C, C++ czy Asembler. Wykorzystanie tych metod jest niezbędne do lepszej optymalizacji działania programu i redukcji zużycia zasobów pamięci oraz procesora.

Uruchomienia algorytmów po stronie języka Java można dokonać za pomocą specjalnie zdefiniowanych funkcji w plikach aplikacji po stronie języka C++. W języku Java należy zdefiniować identyczne funkcje, z uwzględnieniem składni języka programowania jak w języku C++. Pozwala to na połączenie obydwu platform.

Istotnym elementem aplikacji jest również jej wygląd, dlatego zdefiniowano prostą wizualną strukturę aplikacji, której celem jest zapewnienie identycznego układu elementów dla różnych efektów gitarowych (rys. 5). Aplikacja ma liniowy układ graficzny, podzielony na kilka niezależnych od siebie układów relatywnych.

Górna część interfejsu graficznego aplikacji GuitarDroid jest widokiem stałym i ma jedynie przyciski służące do zmiany widoków efektów gitarowych. Druga część widoku została podzielona za pomocą elementu ViewFlipper na niezależne układy i jest zmieniana w zależności od wyboru efektu. Wyboru algorytmu przetwarzania dokonuje się przez dotknięcie przycisku z odpowiednią nazwą danego efektu.

Suwaki służą do zmiany parametrów wybranego efektu. Liczba suwaków zależy od liczby zmiennych parametrów algorytmu przetwarzania. Co najistotniejsze, parametry zmieniane są w czasie rzeczywistym. W polach tekstowych po prawej stronie suwaków ukazane są wartości suwaka i nazwa zmienianego parametru [4].



Rys. 5. Prosty interfejs graficzny aplikacji

5. REALIZACJA PRZEJŚCIÓWKI

Podłączenie gitary do telefonu wymaga zastosowania przejściówki w celu dopasowania impedancyjnego. Wynika to z faktu standaryzacji złącz minijack w smartfonach, które, aby odczytać sygnał wejściowy, wymagają obciążenia trzeciego pinu rezystancją 5 k Ω . W przeciwnym razie w telefonie opcja ta zostaje zablokowana i możliwe jest jedynie odtwarzanie dźwięku.

Aby odseparować od siebie układy elektroniczne gitary i telefonu, zastosowany został tranzystor JFET w konfiguracji wspólnego źródła. Tak skonstruowany układ odgrywa rolę bufora, który jednocześnie wzmacnia sygnał dwukrotnie. Na wyjściu znajduje się potencjometr, działający jako tłumik. Tranzystor w takiej konfiguracji charakteryzuje się – poza dużym wzmocnieniem – odwracaniem fazy, jednak ze względu na opóźnienie, jakie wprowadza Android, ma to niewielkie znaczenie dla brzmienia, więc aspekt ten został pominięty przy dalszym projektowaniu. Całość jest zasilana baterią 9 V [8, 9].

6. TESTY GOTOWEGO ROZWIĄZANIA

Pierwsze testy aplikacji wykonane zostały przy użyciu fabrycznie dołączonego zestawu słuchawkowego wyposażonego w mikrofon. Po odsłuchaniu na słuchawkach efektów zarejestrowanych mikrofonem dźwięków gitary bądź ludzkiego głosu można było stwierdzić percepcyjnie prawidłowość działania zaprojektowanych algorytmów. Punktem odniesienia były próbki dźwiękowe przetworzone profesjonalnymi, programowymi efektami zaimplementowanymi w oprogramowaniu Adobe Audition CS6. Na podstawie subiektywnych

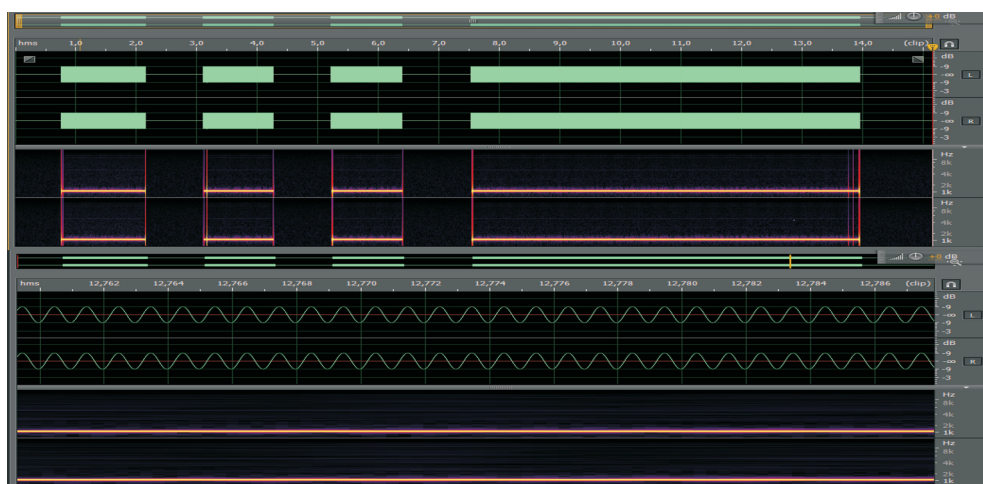
odsluchów zarówno twórców projektów, jak i osób postronnych można było stwierdzić, czy algorytm działa prawidłowo (minimalnie 90% opinii pozytywnych), czy wymaga poprawy. Ostatni etap testów dotyczył zestawionego gotowego układu przejściówki z telefonem.

Testy zostały przeprowadzone za pomocą smartfona Meizu M2 Note opartego na Androidzie w wersji Lolipop 5.1. Za pomocą przejściówki do telefonu przesyłany był sygnał z wyjścia liniowego w karcie dźwiękowej Audiotrack Prodigy Cube, a następnie przetworzony przez aplikację dźwięk trafiał z powrotem do interfejsu.

Dodatkowo wykonane zostały testy subiektywne na smartfonach Sony XPERIA Z2 oraz Samsung Galaxy Tab 2. Podczas testowania słysząc było różnice w opóźnieniu pomiędzy urządzeniami różnej klasy.

7. TESTY EFEKTÓW

Jako sygnał testowy użyto przebiegu sinusoidalnego o częstotliwości 1 kHz i amplitudzie -9 dB odpowiadającej 0 dB w systemach analogowych, przedstawionego na rysunku 6.

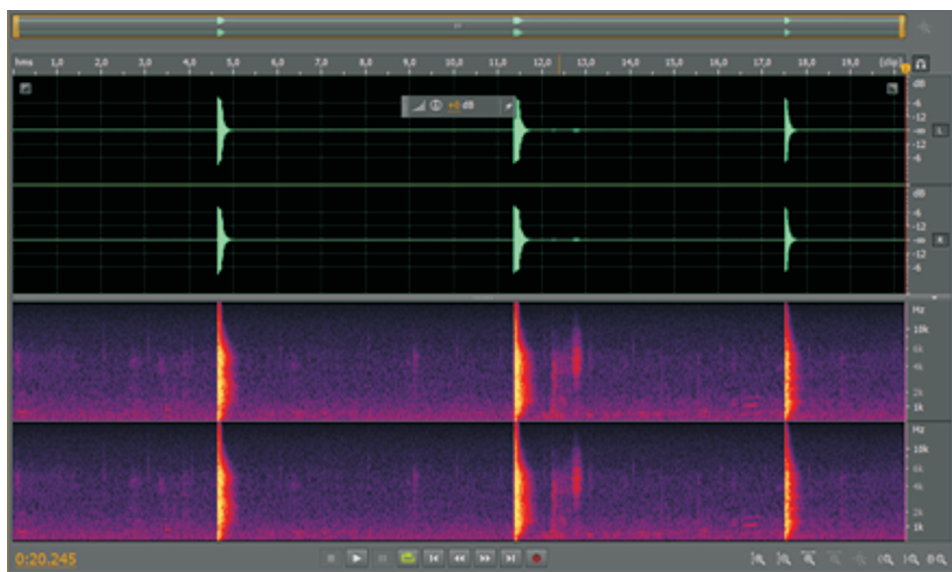


Rys. 6. Sygnał testowy, dolny wykres reprezentuje zbliżenie na obwiednię i widmo

Pierwszym z przetestowanych efektów wbudowanych w procesor było echo. Testowano efekt o zadanych wartościach parametrów, gdzie parametr mix oznacza stosunek sygnału wejściowego (suchego) do sygnału po przejściu przez procesor efektowy (dry), Feedback oznacza sprzężenie zwrotne, a Time czas pogłosu:

- Mix: 0,5,
- Feedback: 0,7,
- Time: 1 s.

Aby prawidłowo sprawdzić jego działanie, układ został pobudzony krótkim impulsem, dzięki czemu w sygnale wyraźnie widać kolejne odbicia. Możliwość zaznaczenia fragmentu i sprawdzenia czasu jego trwania pozwala stwierdzić zgodność działania efektu z konfiguracją, czyli że echa sygnału wejściowego pojawiają się dokładnie w odstępach sekundy. Na przebiegu widmowym widać również, że moc każdego kolejnego sygnału jest mniejsza od poprzedniego. Efekt działa zatem prawidłowo. Przebieg sygnału i wykres widma przedstawia rysunek 7.



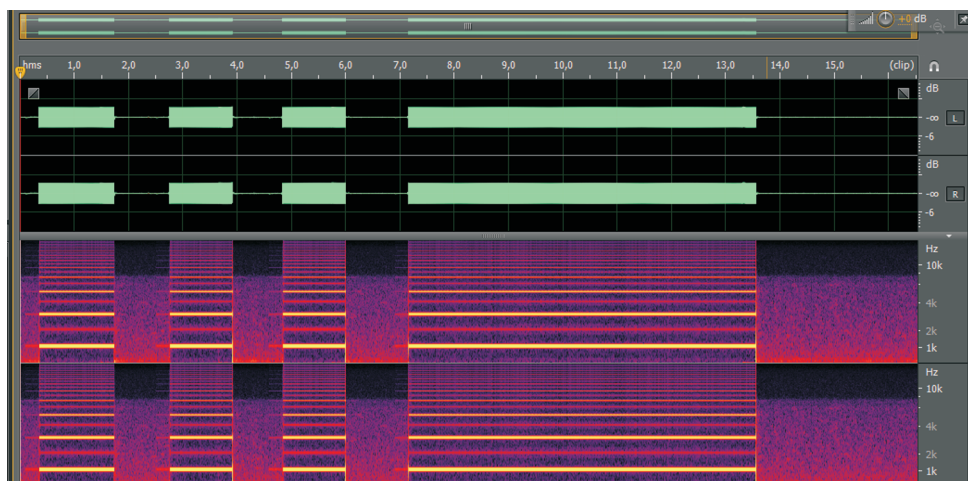
Rys. 7. Analiza amplitudy i widma sygnału po przejściu przez efekt echa

Warto zwrócić uwagę na fakt wprowadzenia do filtracji niestabilności wynikającej z charakterystyki filtru IIR, która tworzy się wraz ze wzrostem wartości parametru feedback. Gdy osiąga ona poziom 0,8, wtedy każdy głośniejszy sygnał powoduje, że na wyjściu pojawia się delikatny świergot.

Kolejnym z testowanych efektów był przester o zadanych parametrach, gdzie Gain odpowiada za wzmocnienie przesterowania w sygnale wyjściowym, a Bias jest parametrem odpowiadający w symulacji prądowi wstecznemu, który nadaje specyficzny ton:

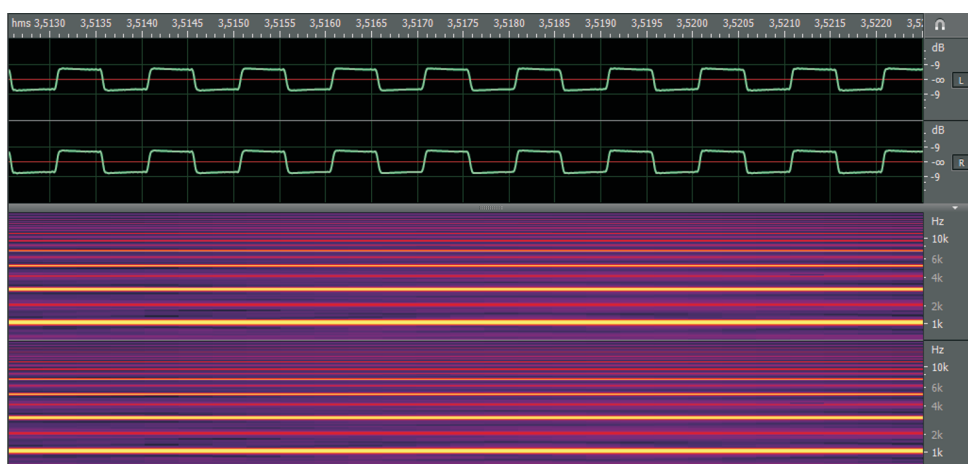
- Gain: 30,
- Bias: 0,5.

Do sprawdzenia poprawności działania efektu wykorzystano cztery sygnały – trzy krótkie i jeden długi, aby móc w pełni pokazać stopień ich zniekształcenia. Jak widać na rysunku 8, na którym przedstawiono widmo i amplitudę otrzymane po przejściu sygnału testowego przez efekt, są one w pełni niezależne od długości trwania sygnału.



Rys. 8. Analiza amplitudy i widma sygnału po przejściu przez efekt przesteru

Na zbliżeniu przebiegu sygnału (rys. 9) można zauważyć, że sygnał sinusoidalny został znacznie zniekształcony, maksima zostały wyrównane, a sygnał przypomina przebieg prostokątny. Taki obraz wynika z samego działania algorytmu, który w początkowej fazie dodaje minimalną składową stałą określoną w projekcie jako Bias, następnie wzmacnia sygnał, mnożąc go przez parametr Gain, aby na końcu dzięki zastosowaniu funkcji tangens hiperboliczny odciąć szczyty sygnału. W wyniku tych zabiegów amplituda sygnału nie zmienia się, ale znacząco zmienia się jego kształt, generując charakterystyczne metaliczne brzmienie.

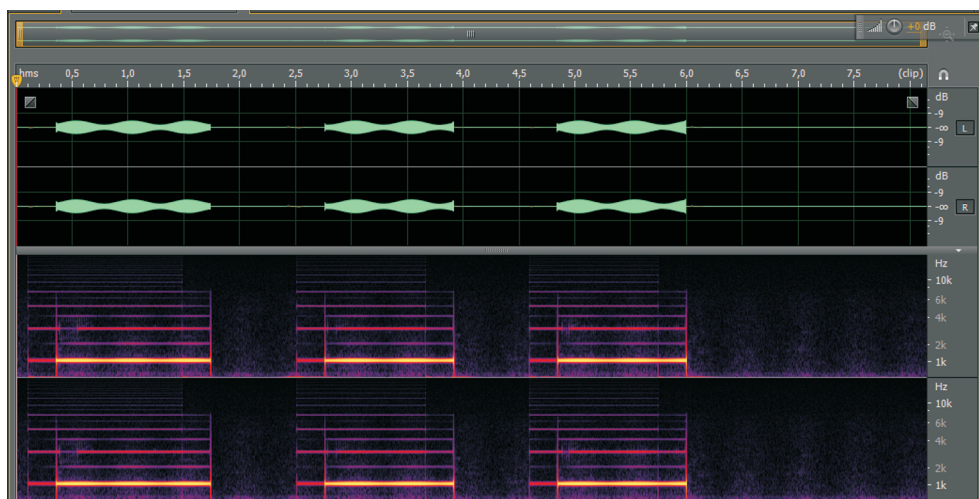


Rys. 9. Zbliżenie na widmo i amplitudę sygnału z nałożonym efektem przesteru

Następnym z zaimplementowanych efektów było tremolo, testowane przy użyciu poniższych parametrów, gdzie Depth oznacza głębokość modulacji, a Frequency częstotliwość, z jaką sygnał jest modulowany:

- Depth: 0,9,
- Frequency: 1 Hz.

Prawidłowość działania efektu sprawdzono na podstawie trzech krótkich sygnałów. Wyniki analizy przedstawiono na rysunku 10.



Rys. 10. Analiza amplitudy i widma sygnału po przejściu przez efekt tremolo

Wyraźnie widać, że amplituda sygnału została ograniczona funkcją sinusoidalną o okresie wynoszącym 1 s (czyli o częstotliwości 1 Hz). Ustawienie wysokiej wartości parametru Depth spowodowało znaczne obniżenie amplitudy w miejscach maksimów funkcji ograniczającej. Takie ustawienie parametrów testowych pomogło w prawidłowej analizie efektu, dzięki której można potwierdzić zgodność pracy zaimplementowanego algorytmu z założeniami teoretycznymi dotyczącymi efektu tremolo.

8. PODSUMOWANIE

Wykorzystane biblioteki są niezwykle rozbudowane i dają olbrzymie możliwości już w swojej podstawowej funkcjonalności, jednak w stosunkowo prosty sposób, wzorując się na już istniejących, można dodawać własne algorytmy. Cały proces przetwarzania i ich łączenia jest już zaimplementowany. JNI jest interfejsem niezbędnym przy współczesnym programowaniu. Korzyści wynikające z jego zastosowania mogą znaleźć odzwierciedlenie

nie tylko w przetwarzaniu sygnałów fonicznych, ale też w dowolnej aplikacji. Wynika to z faktu znacznie szybszej i stabilniejszej pracy aplikacji pisanych w językach C czy C++.

Na podstawie testów można zauważyć, że system Android nie jest dobrze zoptymalizowany pod względem przetwarzania dźwięków. Jego największą wadą w tej dziedzinie jest znaczne opóźnienie, które wprowadza do toru audio. W przypadku użytych bibliotek jest ono rzędu 40 ms w zależności od wykorzystywanego urządzenia. Jest to opóźnienie zbyt duże, aby było akceptowalne do komfortowego grania na gitarze.

Google jako twórca oprogramowania Android z każdą kolejną wersją natywnie wprowadza coraz to lepszą obsługę strumienia audio, np. przez wprowadzenie obsługi *low latency API* (niskoopóźnieniowy interfejs programowania). Jej zastosowanie pozwala obniżyć latencję w kodzie implementowanym w Javie do poziomu ok. 30 ms, co dobrze rokuje na wprowadzenie opisanego produktu na rynek.

LITERATURA

- [1] Szanto G., Vlaskovits P.: Superpowered, <http://superpowered.com> [data dostępu: 20.10.2015]
- [2] Google Inc.: Android Developers, <https://developer.android.com> [data dostępu: 5.2015]
- [3] Orfanidis S.J.: Introduction to Signal Processing, Prentice Hall, Rutgers University, 1995
- [4] Flosisand C.: Float Audio Programing (DSP), <https://christianfloisand.wordpress.com> [data dostępu: 1.10.2015]
- [5] Cronin D.: Examining Audio DSP Algorithms, Dr. Dobb's Journal, 1994, 19 (7), 78–83
- [6] Materiały dydaktyczne z wykładu z realizacji dźwięku, Politechnika Wrocławska, 12.2013
- [7] Ciarkowski A.: DSP++ Library, <https://bitbucket.org/andrzejc/dsp> [data dostępu: 14.08.2015]
- [8] CadSoft Computer: Eagle Easily Applicable Graphical Layout Editor Tutorial Version 7.7.3.0, http://www.cadsoftusa.com/fileadmin/journalist/Documents/V7.3_tutorial_en.pdf [data dostępu: 19.09.2015]
- [9] Google Inc.: Wired Audio Headset Specification, <https://source.android.com/devices/accessories/headset/speification.html> [data dostępu: 26.09.2015]

POLISH HORROR STORY.

HISTORIA

STWORZONA Z EFEKTÓW DŹWIĘKOWYCH –
CZY JEST TO MOŻLIWE?

Polish horror story.

Story created with only sound effects – is it even possible?

Barbara KAPICA, Zuzanna RACIBORSKA

AGH Akademia Górniczo-Hutnicza,
Wydział Inżynierii Mechanicznej i Robotyki,
barbara.kapica@yahoo.com

STRESZCZENIE

W rozdziale opisano projekt, którego celem było stworzenie zróżnicowanych efektów dźwiękowych, które zostały następnie wykorzystane do słuchowiska przedstawiającego krótką historię. Autorki zrealizowały charakterystyczne dla gatunku grozy dźwięki, wykorzystując trzy podstawowe metody.

Pierwszą metodą było nagrywanie dźwięków w odpowiednim miejscu, o różnych porach dnia i w zmiennej sytuacji atmosferycznej (w czasie deszczu czy w czasie silnego wiatru). Ważne było zarejestrowanie jak najwięcej zróżnicowanych efektów dźwiękowych (np. nagranie kroków człowieka na zmiennym podłożu, w różnym obuwiu) i wybranie najodpowiedniejszego nagrania. Drugą metodą uzyskiwania dźwięków była tzw. rekonstrukcja efektów dźwiękowych, np. nagranie dźwięków towarzyszących obieraniu, rozrywaniu i rozgniataaniu owoców i warzyw, następnie zmodyfikowanych za pomocą wtyczek programu DAW. Zastosowano między innymi zmianę prędkości nagrania, zmianę wysokości nagrania, modulacje, filtry górno- i dolnoprzepustowe. Autorki wykorzystały również możliwość nakładania na siebie dźwięków pochodzących z różnych nagrań w celu uzyskania ciekawszego brzmienia. Ostatnią metodą wykorzystaną w badaniu było generowanie dźwięków w programie PureData.

Na podstawie uzyskanego czterominutowego nagrania został przeprowadzony test na małej grupie losowo wybranych osób. Celem było sprawdzenie, czy słuchacz potrafi stworzyć historię wyłącznie na podstawie efektów dźwiękowych (bez bodźca wzrokowego, który zazwyczaj najskuteczniej oddziałuje na odbiorcę). Badanie zostało przeprowadzone drogą internetową, co umożliwiło słuchaczowi przesłuchanie nagrania w komfortowych dla siebie warunkach i uzupełnienie kwestionariusza z pytaniami otwartymi oraz zamkniętymi na temat słuchowiska.

ABSTRACT

The main goal of project described in this chapter was creating diverse sound effects that afterwards were used in a short audio play. Sound design is not about recording (reproducing faithfully the sound), but about creating a new sound which is usually exaggerated for effect. Authors proposed using basic methods to create and reproduce sound, for example, recording at previously chosen localization during different time a day and during different weather conditions. It was very important that the sounds were diverse (recording footsteps with different footwear and different floor). Next method used in this project was recreating sound, which is recording sounds from everyday objects or fruits and vegetables and using basic plugins (for example. time stretching, pitch shifting, low and high pass filters), in order to obtain as realistic sound as it was possible. The last method used in this project was creating sounds using PureData. From recorded sound the story was created, and nextly a survey was conducted in order to test if a listener can reproduce the story that was told in the record (without visual aspect). The participants completed the survey online.

1. WSTĘP

Niewiele osób zdaje sobie sprawę, jak bardzo ważny w filmie jest dźwięk. Przeciętny słuchacz nie słyszy pojedynczych dźwięków, które bardzo mocno oddziałują na emocje widza [1].

Ścieżka dźwiękowa filmu czy słuchowiska składa się zazwyczaj z trzech elementów: efektów dźwiękowych, muzyki oraz dialogów. Każdy z tych elementów jest bardzo ważny i „odpowiada” za nastrój filmu. Jednak widzowie najczęściej koncentrują uwagę na obrazie

oraz dialogach, które wprowadzają historię, dźwięk jest traktowany jako dodatek, który może wzmocnić dany przekaz.

Efekty dźwiękowe w świecie filmowym dzielą się na [2]:

- efekty podstawowe (*hard effects*) – dźwięki takie jak strzały z pistoletu, odgłosy klaksonu czy uderzenia),
- imitacja dźwięku (technika Foley) – dźwięki nagrywane już w postprodukcji w synchronizacji z obrazem (np. kroki),
- dźwięki tła,
- dźwięki elektroniczne,
- dźwięki zaprojektowane – efekty, które są bardzo często niemożliwe do nagrania w rzeczywistości, dlatego najczęściej tworzy się je w programach DAW [2].

Projekt „Polish horror story. Historia stworzona z efektów dźwiękowych – czy jest to możliwe?” składał się z dwóch części. Pierwsza część miała na celu nagranie oraz reprodukcję dźwięków, które stanowiłyby pewnego rodzaju całość. W drugiej części zostało przeprowadzone badanie na grupie osób, których zadaniem było przesłuchanie nagrania oraz udzielenie odpowiedzi na kilka pytań.

Autorki przyjęły hipotezę, że każdy uczestnik badania stworzy swoją historię, niekoniecznie podobną do historii wykreowanej przez autorki. Krótkie „słuchowisko” składało się wyłącznie z efektów dźwiękowych (bez dialogów oraz bez muzyki). W kolejnych częściach rozdziału przedstawiony został proces realizacji poszczególnych efektów dźwiękowych oraz przebieg badania na grupie respondentów. Artykuł kończy analiza uzyskanych wyników.

Przebieg wydarzeń założony przez autorki:

1. Akcja rozgrywa się w starym hangarze (lub stodole), wiatr z ogromną siłą przebija się przez dziurawe drzwi, tworząc ogłuszający hałas; nagle słychać strzał z pistoletu, po którym odlatują wystraszone ptaki (mewy).
2. Następnie można usłyszeć chaotyczne myśli kobiety, która właśnie zastrzeliła swoją ukochaną. Myśli te są niepełnymi zdaniami, często powtarzającymi się, co podkreśla chaos panujący w głowie. Fragment ten miał na celu nakreślenie obrazu (słuchacz może domyśleć się, że doszło do morderstwa, może odczytać rozpacz bohaterki, ale również powód morderstwa – główna bohaterka była wilkołakiem).
3. Rozmowa telefoniczna; osoba dzwoni z pomieszczenia (kuchnia), dzwoni do mężczyzny, który ma pomóc jej pozbyć się ciała. Mężczyzna ten również jest wilkołakiem, jednak słuchacz jeszcze nie jest tego świadomy.
4. Mężczyzna wchodzi do mieszkania, następnie owija folią leżące ciało i przenosi je do kuchni. Zaczyna rozcinać ciało (za pomocą piły mechanicznej), włącza piekarnik oraz odkręca kuchenkę gazową, aby przyrządzić ciało do zjedzenia.
5. Nagle słychać z oddali wycie wilkołaka nawołujące do przemiany.
6. Mężczyzna przemienia się w wilkołaka oraz „pożera” ciało. Na końcu dołącza do stada wilkołaków.

2. REALIZACJA SŁUCHOWISKA

Część nagrań została zrealizowana w Laboratorium Inżynierii Dźwięku Akademii Górniczo-Hutniczej, inne nagrania zostały zrealizowane w mieszkaniu jednej z autorek projektu. Wykorzystano rejestrator Zoom H6 z kapsułami XY lub MS, mikrofon kierunkowy superkardioidalny Rode NTG-2 oraz mikrofon pojemnościowy MXL 770. Zrealizowano wiele takich samych nagrań, jednak z różnymi ustawieniami mikrofonu (mikrofon ustawiony bardzo blisko lub daleko), aby uzyskać jak najwięcej materiału, z którego można było wybrać najodpowiedniejszą próbkę.

2.1. DŹWIĘKI TŁA

Dźwięki tła zostały nagrane w różnych miejscach o różnych porach dnia. Pierwsza scena rozgrywa się w starym budynku nad morzem. Dwa nagrania do niej zrealizowano w Kołobrzegu w hali stoczni. Pierwsze z nich zostało zrealizowane w wietrznym dniu. Wiatr przebiegał przez uszkodzone przez sztorm, dziurawe, blaszane drzwi, wywołując przerażający hałas. Kolejny dzień nagrań był bezwietrzny, jednak deszczowy. Wybranie odpowiednich fragmentów pozwoliło na wprowadzenie nastroju grozy od samego początku nagrania sceny.

Kolejne dźwięki, jakie można usłyszeć, to odgłosy zwierząt. Najpierw krzyk morskich ptaków (mew), następnie odgłosy kota.

Innymi dźwiękami wykorzystanymi w tle były odgłosy dochodzące z mieszkania, w szczególności z kuchni. Do tych dźwięków należały między innymi dźwięk włączania kuchenki gazowej czy pracującego piekarnika.

2.2. EFEKTY DŹWIĘKOWE UZYSKANE Z NAGRAŃ

Kroki stały się bardzo istotnym elementem ścieżki dźwiękowej. Były nagrywane w różnych miejscach w Krakowie, takich jak: pokój ze starą, skrzypiącą, drewnianą podłogą; kuchnia, w której podłoga pokryta była kafelkami, oraz klatka schodowa. Dźwięk skrzypiących drzwi wykorzystany w nagraniu był również rzeczywistym brzmieniem starych drewnianych drzwi, znajdujących się w jednej z przedwojennych krakowskich kamienic.

2.3. PROJEKTOWANIE DŹWIĘKU Z WYKORZYSTANIEM PROGRAMU PUREDATA

Dźwięki wybierania numeru telefonu, sygnału oczekującego na połączenie oraz dźwięk sygnału zajętej linii zostały stworzone w programie PureData. Dźwięki wybierania numeru telefonicznego operują na sygnalizacji tonowej DTMF (*Dual Tone Multi Frequency*). Poszczególne tony powstają dzięki nałożeniu na siebie dwóch sygnałów sinusoidalnych o odpowiednich częstotliwościach przypisanych do konkretnego przycisku. Dźwięki sygnału połączenia oczekującego oraz sygnału zajętej linii są tzw. specjalnymi

tonami informacyjnymi (*Special Information Tones*, SIT), które charakteryzują się inną częstotliwością sygnału sinusoidalnego od sygnałów DTMF oraz sprecyzowanym czasem trwania. Przy projektowaniu tych dźwięków skorzystano z książki *Sound Design* napisanej przez Andy'ego Farnella [3].

2.4. REKONSTRUKCJA DŹWIĘKÓW (*SOUND DESIGN*)

Dźwięk wystrzału z broni palnej został uzyskany z nagrania dźwięku przebijanego balonu. Nagranie to zostało następnie poddane silnej kompresji (bardzo krótki czas ataku). Próbkę dźwięku poddano również filtracji, aby odpowiednie pasmo częstotliwości było bardziej słyszalne. Na koniec został dodany długi pogłos z „metalicznym” zakończeniem, aby zasugerować odpowiednie miejsce wydarzenia tworzonej historii.

Jednym z najważniejszych elementów ścieżki dźwiękowej były dźwięki rozkrajania ciała za pomocą piły mechanicznej, odgłosy łamanych kości i chrząstek, a także kapiącej oraz rozpryskującej się naokoło krwi.

Efekt dźwiękowy piły mechanicznej uzyskano z dźwięku, jaki wydawała gazeta przyłożona do pracującego koła roweru – rower postawiono na siodełku, jedna osoba kręciła pedałami, natomiast druga przykładła gazetę do poruszającego się tylnego koła. Nagranie zostało poddane korekcie częstotliwości (wyciszono częstotliwości powyżej 8 kHz, które dawały charakterystyczny dźwięk poruszających się pedałów rowerowych. Drugą warstwą w nagraniu był sygnał piłokształtny o częstotliwości 65 Hz, który zapewnił więcej „wibracji” oraz niskich częstotliwości. Miało to na celu wzmocnienie nastroju grozy, utrzymującego się przez całość nagrania.

Do uzyskania kolejnych dźwięków niezbędnych do stworzenia efektu grozy konieczne było wykorzystanie warzyw i owoców[2]. Łamany, zmrożony por naśladował dźwięk łamanych kości, natomiast dźwięk rozdzierania pojedynczych liści kapusty imitował dźwięk łamanych chrząstek. Efekty dźwiękowe towarzyszące rozrywaniu ciała i naśladujące odgłosy wyciekających wnętrzności człowieka uzyskano przez zgniatanie pomidora oraz wyszarpywanie mięszu pomarańczy. Wszystkie dźwięki zostały na siebie odpowiednio warstwowo nałożone, aby efekt był bardziej sugestywny.

Jednym z trudniejszych elementów do zrealizowania podczas nagrania okazało się uzyskanie dźwięku transformacji wilkołaka. Sama transformacja w warstwie dźwiękowej składała się z dźwięków, takich jak:

- przyspieszone bicie serca – wykorzystano nagranie surda, brazylijskiego instrumentu perkusyjnego, przepuszczone przez filtr dolnoprzepustowy [4],
- łamanie pojedynczych kości – różnorodne nagrania łamanego, zmrożonego pora [2],
- wyrastająca sierść – szybkie ocieranie materiałem o rękę,
- warczenie zwierzęcia – wykorzystano chrapanie mężczyzny, którego dźwięk został obniżony o kilka tonów,
- odgłosy wydawane przez człowieka podczas bolesnej przemiany – aktor oddający odgłosy człowieka, który cierpi.

Kolejnym niezwykle istotnym dla całości nagrania dźwiękiem, były odgłosy pożerania ludzkiego ciała przez bestię, czyli wilkołaka. Autorki wykorzystały do tego nagrane wcześniej odgłosy towarzyszące zgniataniu i obieraniu warzyw i owoców, takich jak: kapusta, pomidor, pomarańcza, orzechy oraz por, a także odgłosy wydawane przez człowieka (np. odgłos oddychania). Dźwięki wydobyte ze wspomnianych wcześniej warzyw i owoców zostały na siebie nałożone, aby nadać całemu efektowi głębię [2]. Autorki wykorzystały również funkcję odwrócenia danej ścieżki i podłożenia jej w tej wersji pod oryginalne nagranie. Autorki zmodyfikowały również głośność poszczególnych nagrań. Dźwięki towarzyszące czynnościom gniecenia, rozrywania niektórych warzyw (kapusta, pomidor) zostały w nagraniu uwydatnione, inne dźwięki (związane z czynnościami rozgniatacia pomarańczy czy tłuczenia orzechów) były jedynie tłem.

3. BADANIA ANKIETOWE

Badanie ankietowe zostało zrealizowane drogą internetową, aby uczestnicy w zaciszu swojego mieszkania mogli przesłuchać nagranie w całości i odpowiedzieć na pytania znajdujące się w ankiecie. Ankieta została stworzona w Formularzach Google w wersji polskiej oraz w wersji angielskiej.

Przebieg badania

- uczestnicy brali udział w ankiecie dobrowolnie, zostali poproszeni o przesłuchanie całego nagrania oraz fragmentów przez słuchawki (aby mogli usłyszeć wszystkie szczegóły),
- w pierwszej części ankiety należało podać kilka informacji na temat respondenta (wiek, płeć, narodowość, ulubiony gatunek filmowy),

Następnie uczestnicy zostali poproszeni o przesłuchanie nagrania tylko raz w całości i udzielenie informacji dotyczących:

- miejsca rozgrywanych wydarzeń,
- gatunku filmowego, z którym kojarzy się im nagranie,
- liczby bohaterów,
- bohaterów,
- przebiegu wydarzeń.

W kolejnej części ankiety uczestnicy mogli przesłuchiwać nagranie dowolną liczbę razy, aby odpowiedzieć na bardziej szczegółowe pytania. Zostali poproszeni o:

- interpretację fragmentu z wystrzałem z broni – badani mieli za zadanie odpowiedzieć na pytanie, co sugeruje ten dźwięk,
- ponowne wymienienie miejsca rozgrywanych wydarzeń (w celu sprawdzenia, czy po kilkakrotnym przesłuchaniu respondenci nie zmienili zdania co do miejsca akcji albo czy potrafili dokładniej określić miejsce wydarzeń),

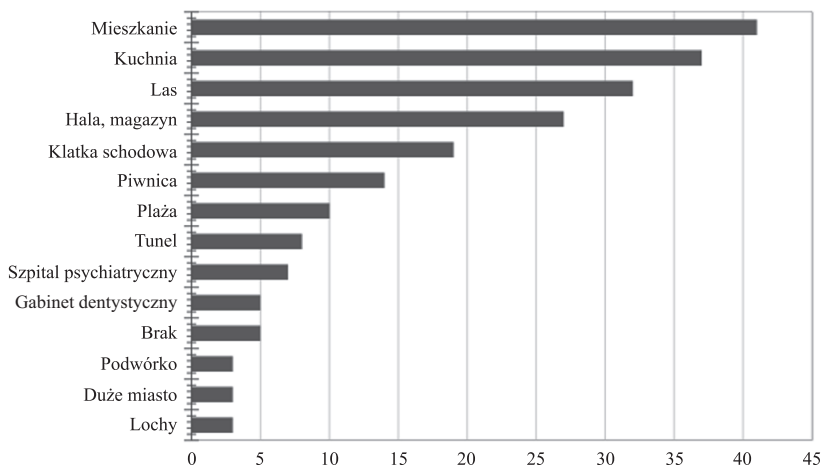
- wymienienie w punktach przebiegu wydarzeń (w celu sprawdzenia, czy po wielokrotnym odsłuchaniu nagrania słuchacz może dokładniej wykreować fabułę),
- interpretację fragmentu z dźwiękiem piły mechanicznej oraz określenie skojarzeń towarzyszących temu dźwiękowi,
- przesłuchanie fragmentu transformacji wilkołaka oraz udzielenie odpowiedzi na pytanie, z jaką istotą nadprzyrodzoną kojarzy się respondentom ten dźwięk,
- stwierdzenie, co według nich dzieje się z istotą nadprzyrodzoną w nagraniu.

Respondenci

W badaniu wzięło udział 59 osób, w tym 49 kobiet, dziewięciu mężczyzn oraz jedna osoba, która odmówiła ujawnienia swojej płci. Średnia wieku respondentów wynosiła 24 lata, najmłodszy uczestnik miał 14 lat, natomiast najstarszy 52 lata. Najwięcej osób było w wieku 21 i 22 lat (razem 31 osób). W ankiecie uczestniczyło 51 osób polskiego pochodzenia oraz osiem osób z zagranicy (jedna osoba pochodzenia latynoskiego, jedna osoba pochodzenia francuskiego oraz sześć osób ze Stanów Zjednoczonych).

4. WYNIKI I ANALIZA

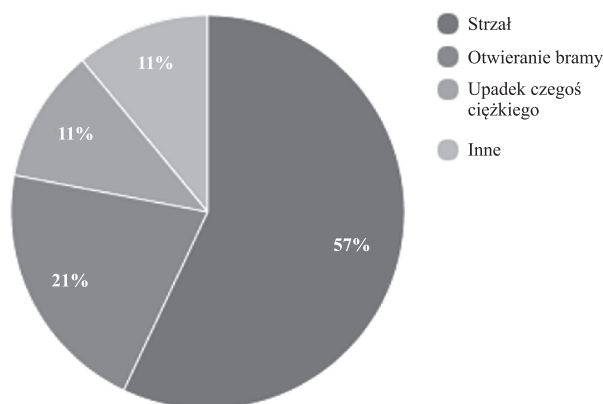
Osoby badane zostały poproszone o określenie miejsca rozgrywanych wydarzeń. Duża różnorodność odpowiedzi była spowodowana rodzajem zadanego pytania w ankiecie. Było to pytanie otwarte, umożliwiające respondentom większą swobodę wypowiedzi i niewpływające na ich interpretację. Rozkład odpowiedzi ankietowanych przedstawiono na rysunku 1.



Rys. 1. Wykres przedstawiający odpowiedzi ankietowanych na pytanie związane z miejscem rozgrywanych wydarzeń

Większość badanych osób określiła miejsce wydarzenia. Jedynie 5% badanych nie udzieliło odpowiedzi (na wykresie zaznaczone, jako „brak”). Najczęściej ankietowani jako miejsce akcji wskazywali mieszkanie, wiele osób dokładniej precyzowało dźwięki jako odgłosy dochodzące z kuchni. Ponad 30% badanych jako miejsce akcji wskazywało las. Najwięcej trudności sprawiła osobom badanym pierwsza scena nagrania. Miejsce, w którym powstało nagranie, była wielka hala. Jedynie nieco ponad 25% zidentyfikowała to miejsce poprawnie. Wiele osób wskazało na piwnicę (14%), tunel (8%), podwórko 3%. Ze względu na pojawienie się odgłosów mew wiele osób jako miejsce akcji wskazało teren nad morzem czy plażę (10%). Na lochy wskazało 3% respondentów, na szpital psychiatryczny 7%, a na gabinet dentystyczny 5%.

Drugie pytanie otwarte dotyczyło jednego z pierwszych wyrazistych dźwięków w nagraniu, czyli wystrzału z broni palnej. Wyniki zostały przedstawione na rysunku 2. Ponad połowa osób badanych (57%) zinterpretowała dźwięk jako wystrzał z broni, 21% uważało ten dźwięk za otwieranie bramy, 11% nie potrafiło dokładnie sprecyzować i często określało dźwięk jako „upadek czegoś ciężkiego”,

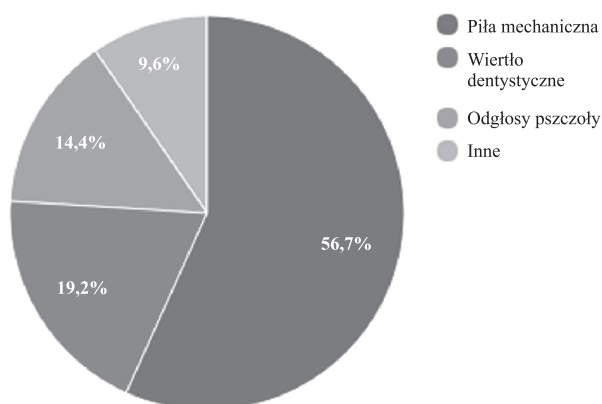


Rys. 2. Wykres przedstawiający odpowiedzi ankietowanych na pytanie o dźwięk wystrzału z pistoletu

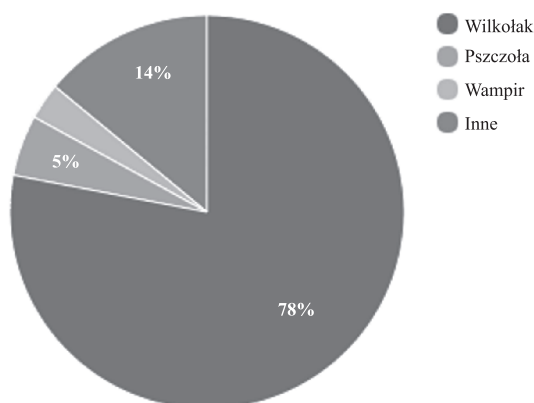
Kolejne pytanie otwarte dotyczyło dźwięku wydawanego przez piłę mechaniczną. Wyniki przedstawiono na rysunku 3.

Zdecydowana większość (56,7%) poprawnie zdefiniowała dźwięk. Z kolei 12 osób (19,2%) uznało ten dźwięk za odgłos wiertła dentystycznego. Zaskakującą odpowiedzią, która pojawiła się kilkukrotnie, była interpretacja dźwięku jako odgłosu pszczoły (14,4%). Pozostałe 9,6% ankietowanych nie potrafiło zinterpretować usłyszanego dźwięku.

Następnie osoby badane zostały poproszone o zinterpretowanie dźwięku przemiany męczyzny w wilkołaka.

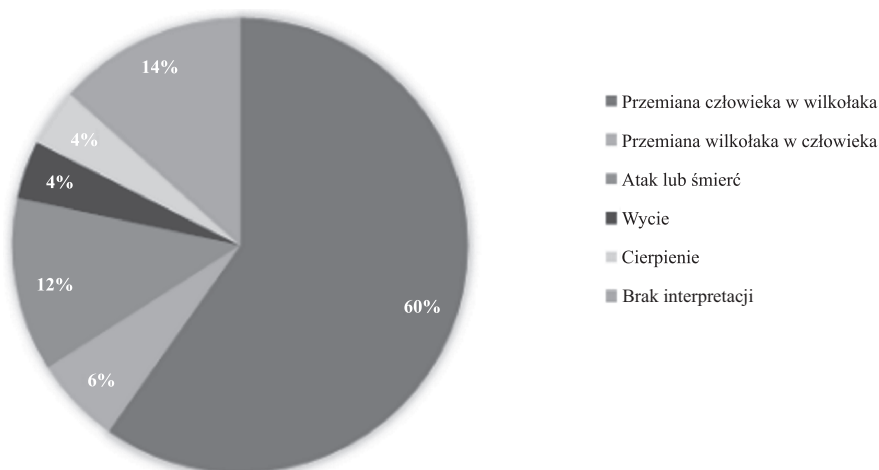


Rys. 3. Wykres przedstawiający odpowiedzi ankietowanych na pytanie o dźwięk piły mechanicznej



Rys. 4. Wykres przedstawiający odpowiedzi ankietowanych na pytanie o istotę nadprzyrodzoną, z którą kojarzy się podany dźwięk

Wyniki zostały przedstawione na rysunku 4 (rozpoznanie istoty nadprzyrodzonej) oraz na rysunku 5 (identyfikacja zjawisk, którym podlega, i uczuć, które temu towarzyszą). Aż 78% uznało odgłosy istoty nadprzyrodzonej za odgłosy wilkołaka, natomiast zaskakująco 60% badanych zinterpretowało dźwięki jako przemianę mężczyzny w wilkołaka (dodatkowo 6% błędnie uznało transformację wilkołaka w mężczyznę). Usłyszanych dźwięków nie potrafiło zinterpretować 14% uczestników ankiety. Pojawiło się też kilka odpowiedzi, że istotą nadprzyrodzoną w słuchowisku może być zombie lub wampir. Kolejną zaskakującą odpowiedzią było podanie przez kilku uczestników ankiety (5%) zmutowanej pszczoły.

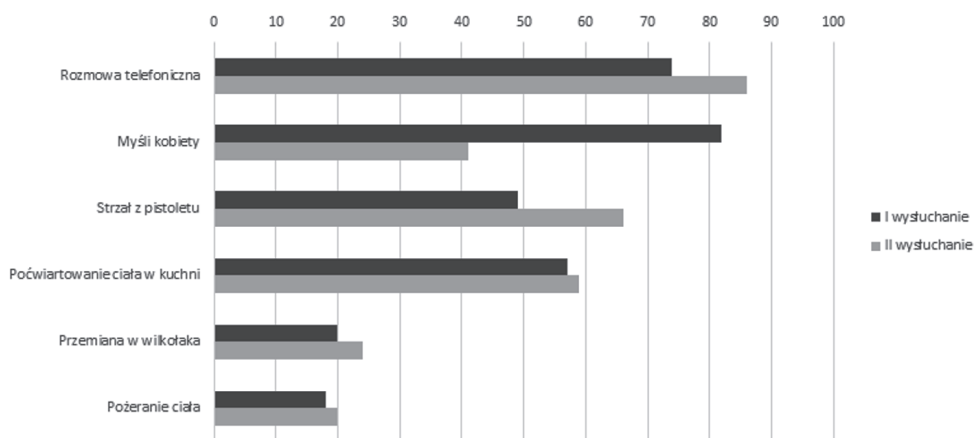


Rys. 5. Wykres przedstawiający odpowiedzi ankietowanych na pytanie, co dzieje się z nadprzyrodzoną istotą

Najważniejszym celem badania było sprawdzenie, czy osoby potrafią określić wydarzenia oraz czy historia stworzona przez respondentów jest zgodna z historią założoną przez autorki. Pytania były otwarte, dlatego przy określeniu zgodności brano pod uwagę jedynie główne elementy, takie jak:

- wystrzał z pistoletu,
- myśli kobiety,
- rozmowa telefoniczna,
- poćwiartowanie ciała w kuchni,
- przemiana w wilkołaka,
- pożeranie ciała.

Rozpoznanie poszczególnych elementów nagrania podczas pierwszego oraz drugiego odsłuchania zostało przedstawione na rysunku 6. Po pierwszym wysłuchaniu nagrania 75% osób badanych potrafiło zidentyfikować rozmowę telefoniczną. Po kolejnym przesłuchaniu wartość wzrosła do 85%. Zaskakująco ponad 80% zidentyfikowało w pierwszym słuchaniu fragment z myślami kobiety, niestety po kolejnym przesłuchaniu wymienilo go jedynie 40%. Strzał z pistoletu został zinterpretowany lepiej w kolejnym przesłuchaniu (66%). Osoby badane największy problem miały z interpretacją ostatniej części „słuchowiska”, jedynie około 20% zidentyfikowało pożeranie ciała i przemianę w wilkołaka. Wiele osób nie zinterpretowało dźwięku wydawanego przez istotę nadprzyrodzoną i traktowało końcowy fragment jako morderstwo. Nie ma znaczących różnic w rozpoznawaniu całej historii w pierwszym przesłuchaniu, jak i w drugim.



Rys. 6. Wykres przedstawiający rozpoznanie przez respondentów poszczególnych elementów historii

Następnie wyniki zostały przeanalizowane. Badano, czy respondenci na podstawie wysłuchanych dźwięków potrafili stworzyć fabułę założoną przez autorki. W tym celu oceniono liczbę poprawnie wymienionych przez respondentów rozgrywających się wydarzeń. Aż 12% nie zinterpretowało żadnego wydarzenia zgodnie z założeniami autorek. Tyle samo osób rozpoznało tylko jeden z elementów historii, w większości była to rozmowa telefoniczna. Najwięcej osób ankietowanych (29%) potrafiło zinterpretować trzy wydarzenia. Żadna z osób nie wymieniła wszystkich sześciu wydarzeń.

5. PODSUMOWANIE I WNIOSKI

Celem badania było sprawdzenie, czy osoby ankietowane będą potrafiły ułożyć pewną historię na podstawie wysłuchanych dźwięków. Okazuje się, że zaledwie 3% badanych (dwie osoby) nie poradziły sobie z tym zadaniem i nie były w stanie przedstawić jakiejkolwiek opowieści. Pozostałe osoby interpretowały przesłuchane nagranie różnie, choć już po pierwszym przesłuchaniu zdecydowana większość stwierdziła, że słuchane nagranie jest z gatunku horror czy thriller.

Większość osób badanych na swój sposób interpretowała słuchowisko i przedstawiała historię, dopasowując poszczególne dźwięki do różnych elementów. Jedną z najbardziej zaskakujących odpowiedzi było zinterpretowanie dźwięku piły mechanicznej jako dźwięku zmutowanej pszczoły. Pojawiło się aż 14% takich odpowiedzi. Co więcej, osoby, które w ten sposób określiły ten dźwięk, wskazały pszczołę jako głównego sprawcę morderstwa. W ponad 50% odpowiedzi pojawił się wątek morderstwa i cięcia zwłok. Jednak tylko 20% ankietowanych odnalazło w historii wątek pożerania ciała, który dla autorek był oczywisty.

Odpowiadając na pytanie zawarte w tytule projektu – okazuje się, że można opowiedzieć historię tylko na podstawie efektów dźwiękowych. W przypadku pozbawienia człowieka bodźców wzrokowych, trzeba mu zapewnić bardzo wyraziste i oczywiste bodźce słuchowe, aby mógł poprawnie odczytać intencje autora.

LITERATURA

- [1] Przedpeńska-Bieniek M.: Dźwięk w filmie, Warszawa 2009
- [2] Viers R.: The Sound Effects Bible, 3rd edition, Studio City, Los Angeles, CA, USA 2008
- [3] Farnell A.: Designing Sound, Cambridge, MA, USA 2010
- [4] <http://www.epicsound.com/sfx> [data dostępu: 18.02.2016]

NIECH ŻYJE TAŚMA!

Tape's not dead!

Mateusz WIDOMSKI

AGH Akademia Górniczo-Hutnicza,
Wydział Inżynierii Mechanicznej i Robotyki,
mateuszwidomski@gmail.com

STRESZCZENIE

Praca prezentuje zasady działania urządzeń wykorzystywanych do zapisu oraz odtwarzania dźwięku z nośników magnetycznych. Przedstawione zostały wady i zalety analogowych, magnetycznych nośników dźwięku. Zaprezentowane zostały rozwiązania konstrukcyjne analogowych procesorów dźwięku wykorzystujących taśmy magnetyczne. Przybliżona została historia oraz techniki produkcji dźwięku z użyciem nośników magnetycznych.

ABSTRACT

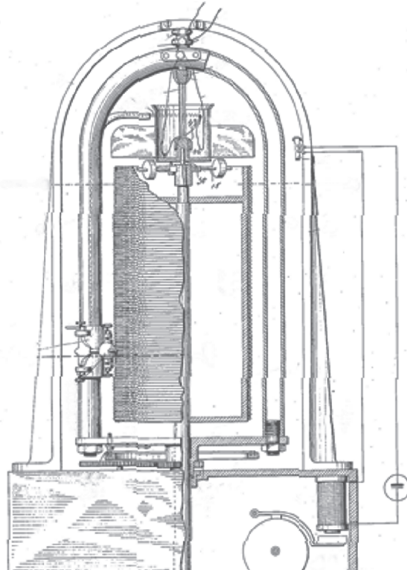
The paper begins with the description of the technique of magnetic tape sound recording, equipment used for recording and reproduction of sound stored on magnetic media. It presents the advantages and disadvantages of analog magnetic sound carriers and their causes. Design solutions of analog sound processors that use magnetic tape have been approximated. The article presents the history and techniques of sound production with the use of magnetic media.

1. HISTORIA MAGNETYCZNEGO ZAPISU DŹWIĘKU

Początki magnetycznego zapisu dźwięku sięgają XIX wieku. Pierwszą osobą, która opisała ten proces, był Oberlin Smith [1], który 8 sierpnia 1888 roku opublikował artykuł opisujący możliwość magnetycznego zapisu dźwięku na nitce z wplecionymi opiłkami metalu. Drobinę metalu miały być namagnesowane proporcjonalnie do zmiennego prądu pochodzącego z mikrofonu. Nie zbudował on jednak urządzenia wykorzystującego tę zasadę do nagrywania dźwięku.

Konstrukтором pierwszego działającego urządzenia nagrywającego dźwięk na nośniku magnetycznym był Valdemar Poulsen [2]. Jego urządzenie miało jedną głowicę magnetyczną, która stosowana była do zapisu i odczytu dźwięku. Zapis odbywał się na metalowym drucie nawiniętym na cylinder (rys. 1). Urządzenie nie miało wzmacniaczy akustycznych, sygnał dostarczony przez mikrofon był na tyle silny, że nagrany dźwięk można było odtwarzać na słuchawkach podłączonych bezpośrednio do głowicy.

Wynalazcą taśmy magnetycznej był Fritz Pfleumer [3], który był także konstruktorem pierwszego magnetofonu szpulowego (rys. 2). Był on ekspertem w sprawach papierów specjalistycznych, udało mu się opracować proces produkcji metalizowanych brązem złotych gilz papierosowych. Wykorzystując ten proces i prace Poulsena [4], Pfleumer wykonał taśmę magnetyczną, w której zmienił brąz na tlenek żelaza będący materiałem ferromagnetycznym. Swoją produkt nazwał „dźwiękowym papierem”, patent na niego otrzymał 1 stycznia 1928 roku. Jakość nagrań wykonanych na papierowej taśmie była niższa niż na stalowym drucie lub taśmie. Taśma papierowa miała jednak zdecydowaną przewagę nad nośnikami metalowymi. Pozwalała na łatwą edycję nagrań za pomocą nożyczek i kleju, a nie obcęgow i lutownicy, które były niezbędne do produkcji dźwiękowej w przypadku starszych nośników magnetycznych. Pfleumer, zdając sobie sprawę z ograniczenia możliwości rozwoju swojego patentu, 28 listopada 1932 roku podpisał kontrakt z niemieckim Allgemeine Elektrizitäts-Gesellschaft (AEG).



Rys. 1. Szkic telegrafonu Poulsena z jego patentu [4]



Rys. 2. Fritz Pfelemer ze swoim magnetofonem

Źródło: https://regmedia.co.uk/2013/04/12/fritz_pfleumer_with_tape_recorder.jpg [data dostępu: 21.02.2016]

Zespół AEG wprowadził zmiany w systemie Pfeumera. W 1933 roku Eduard Schuller opracował okrągłe głowice ze szczeliną w rdzeniu, które zastąpiły mniej sprawne proste głowice. W 1934 roku zespół zmienił także nośnik materiału ferromagnetycznego z papieru na celulozę. Spowodowane to było małą wytrzymałością papieru na naciąganie, zrywaniem się oraz zmianami wymiarów wywołanymi przez wilgoć. Od momentu wykorzystania celulozy rozpoczęła się współpraca AEG z Badische Anilin- und Soda-Fabrik (BASF).

Od tej chwili AEG zajęło się udoskonalaniem sprzętu do nagrywania i odtwarzania taśm magnetycznych, a Badische Anilin- und Soda-Fabrik unowocześnianiem nośników.

Pierwszym praktycznym i komercyjnym urządzeniem odtwarzającym taśmy magnetyczne, produkowanym przez AEG, był Magnetophon K1 zaprezentowany w 1935 roku na wystawie Berlin Radio Show. Rozwinięto też technikę nagrań, eliminując zniekształcenia między innymi przez dodanie do nagrywanego sygnału niewielkiej składowej zmiennej o dużej, ultradźwiękowej częstotliwości.

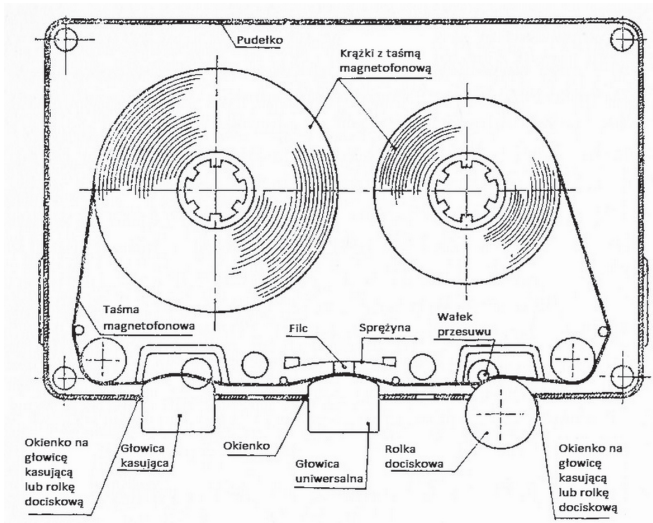
Badania Niemców pozostawały tajne do zakończenia drugiej wojny światowej. Po ich odtajnieniu nastąpił szybki rozwój technik magnetycznego zapisu dźwięku i danych, co doprowadziło do powstania kaset magnetofonowych zaprezentowanych przez firmę Philips w 1962 roku. Dalszy rozwój elektroniki pozwolił na zbudowanie przenośnych odtwarzaczy kaset nazywanych popularnie walkmanami od nazwy produktu firmy Sony.

2. TECHNIKA MAGNETYCZNEGO ZAPISU I ODCZYTU DŹWIĘKU

Sygnał zapisywany jest na taśmie za pomocą elektromagnetycznej głowicy o okrągłym rdzeniu ze szczeliną umieszczoną nad taśmą (rys. 3). Ta sama głowica może być używana także do odtwarzania dźwięku z taśmy po jego nagraniu, głowica taka nazywana jest uniwersalną. Nagranie wykonywane jest na taśmie całkowicie rozmagnesowanej. Rozmagnesowanie wykonywane jest przez głowicę kasującą, która ma większą szczelinę niż głowica nagrywająca. Przez uzwojenia głowicy kasującej przepływa prąd o dużej, ultradźwiękowej częstotliwości. Duża szczelina powietrzna głowicy i duża częstotliwość sygnału doprowadzonego do niej powodują wielokrotne namagnesowanie każdego fragmentu taśmy w obu kierunkach, co powoduje rozmagnesowanie taśmy [5].

Odczyt sygnału z taśmy odbywa się za pomocą głowicy odtwarzającej lub uniwersalnej. Polega on na indukowaniu się napięcia w zwojach głowicy. Efekt ten jest spowodowany przeniesieniem się zmiennego pola magnetycznego taśmy poruszającej się pod szczeliną powietrzną głowicy na jej rdzeń. Wyindukowany sygnał jest następnie wzmacniany i kondycjonowany, co pozwala na dalsze jego przetwarzanie lub odtwarzanie.

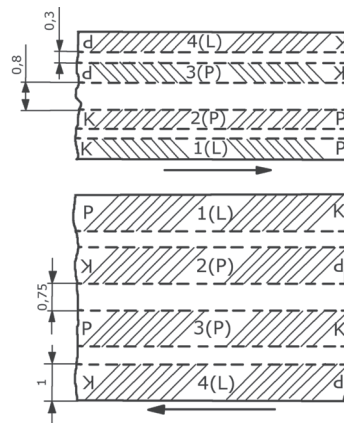
Prędkość liniowa taśmy pod głowicami podczas odtwarzania i nagrywania jest stała i znormalizowana. Pozwala to na dokładne odwzorowanie sygnału pierwotnego. Taśma napędzana jest przez wałek przesuwu sztywno połączony z silnikiem napędowym. Jednocześnie taśma musi być stale odbierana przez prawą szpulę lub lewą w systemach autorewers odtwarzających kasety w obie strony. Prędkość obrotowa szpuli jest zmienna i zależy od promienia szpuli razem z nawiniętą już taśmą. Odpowiednia prędkość szpuli odbierającej taśmę uzyskiwana jest przez zastosowanie sprzęgła ze ślizgiem w jej napędzie. Sprzęgło ślizga się, zmniejszając prędkość obrotową szpuli i synchronizując jej prędkość liniową z prędkością liniową taśmy pod głowicami. W ten sposób taśma jest równomiernie zwijana na szpuli odbierającej, jednocześnie będąc odtwarzana ze stałą prędkością [6].



Rys. 3. Rysunek kasety magnetofonowej wraz z głowicami i napędem [5]

Zapis na taśmach wykonywany jest najczęściej na dwóch ścieżkach dla nagrań monofonicznych lub na czterech dla nagrań stereofonicznych (rys. 4).

Zapis mono odbywał się na ścieżkach 1 i 2 oraz 3 i 4, co nie pozwalało na odtwarzanie stereofonicznych taśm 6,80 mm w magnetofonach monofonicznych i odwrotnie, taśm monofonicznych w magnetofonach stereofonicznych, ponieważ głowica monofoniczna obejmowała ścieżki 1 i 2 lub 3 i 4, a podwójna stereofoniczna 1 i 3 lub 2 i 4 (rys. 4).



Rys. 4. Położenie i szerokość ścieżek od strony głowicy na taśmach 3,81 mm (kasetowa) oraz 6,80 mm (szpulowa) w zapisach czterościeżkowych. P – początek ścieżki dźwiękowej, K – koniec ścieżki dźwiękowej. Oznaczenie L i P oznaczają ścieżki lewą i prawą [5]

3. WADY I ZALETY MAGNETYCZNYCH SYSTEMÓW DŹWIĘKOWYCH

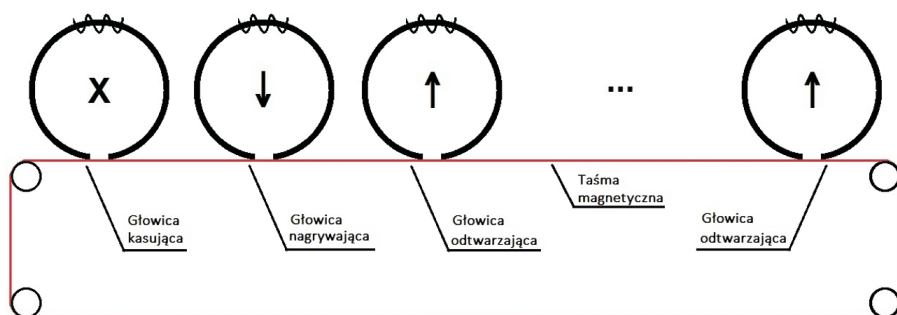
Największą zaletą nośników magnetycznych jest prostota produkcji i reprodukcji dźwięku wykorzystujących taśmy magnetyczne. Do dzielenia i łączenia ścieżek wystarczają nożyczki i taśma klejąca. Do skopiowania taśm potrzebne są dwa magnetofony szpulowe, w których sygnał wyjściowy odtwarzającego podłączony jest jako sygnał wejściowy nagrywającego, lub potrzebny jest magnetofon kasetowy z dwoma szufladami na kasety. Wciąż istnieją, a także powstają nowe studia nagraniowe, które używają tylko systemów analogowych wykorzystujących taśmy jako nośniki dźwięku. Doświadczony inżynier dźwięku może pracować z taśmą równie szybko jak z programem komputerowym DAW (*digital audio workstation* – stacja robocza audio). Jednak systemy analogowe wymagają osobnych, elektronicznych procesorów dźwięku, a nie procesorów cyfrowych w programach DAW. Wymagają one odtworzenia całego sygnału i powtórnego nagrania go po przetworzeniu, natomiast procesory cyfrowe działają automatycznie, przetwarzając cały sygnał w bardzo krótkim czasie.

Możliwość kopiowania rozpowszechniła nośniki magnetyczne w latach osiemdziesiątych i dziewięćdziesiątych ubiegłego wieku. Możliwe było wykonywanie wielu kopii w krótkim czasie oraz tworzenie własnych składanek. Taśmy dźwiękowe wyparły wtedy w dużym stopniu płyty winylowe, do których dostęp był ograniczony, a samodzielne kopiowanie prawie niemożliwe. Ponadto nośniki magnetyczne pozwalały na wielokrotny zapis, co nie było możliwe w przypadku produkowanych masowo nośników mechanicznych takich jak płyty winylowe czy walce woskowane oraz nośników współczesnych: płyt CD i DVD.

Dużą wadą taśm magnetycznych jest ich naciąganie się po wielokrotnym odtworzeniu, przez co dźwięk jest zniekształcony. Częstotliwości w naciągniętej części taśmy obniżają się, ponieważ prędkość liniowa odtwarzania jest znormalizowana i nie jest synchronizowana na podstawie długości taśmy. Kolejną wadą nośników magnetycznych jest wpływ uszkodzeń mechanicznych na odtwarzany sygnał. Każde zagięcie taśmy powoduje trzaski i szумы w sygnale dźwiękowym, co uniemożliwia użycie zniszczonego fragmentu taśmy. Podobny efekt dźwiękowy wywoływany jest przez zabrudzoną taśmę lub głowicę, dlatego należy je czyścić. Długotrwałe tarcie głowicy taśmą powoduje namagnesowywanie się głowicy, co w efekcie może prowadzić do zmiany sygnału zapisanego na taśmie. To samo tarcie powoduje ścieranie się głowicy i zmniejszenie jej czułości.

4. ANALOGOWE, TAŚMOWE PROCESORY DŹWIĘKU

Powstało wiele analogowych procesorów dźwięku wykorzystujących taśmy magnetyczne. Jednym z nich jest echo taśmowe. Urządzenie to zawiera bezkończącą taśmę magnetyczną, jedną głowicę kasującą, jedną nagrywającą i wiele odtwarzających (rys. 5).



Rys. 5. Schemat echa taśmowego

Opóźnienie generowanych odbić dźwięku można ustawić, zmieniając odległość między głowicami odtwarzającymi lub prędkość przesuwu taśmy. Nie należy jednak zmniejszać prędkości taśmy do wartości poniżej normatywnych prędkości odtwarzania danego typu taśm, ponieważ może to negatywnie wpłynąć na zakres częstotliwości przenoszonych przez taśmę. Optymalna będzie prędkość odpowiednio duża, którą można będzie zwiększać i zmniejszać do normatywnej wartości.

Kolejnym procesorem dźwięku, który może być używany „na żywo”, jest flanger. Uzyskuje się go za pomocą dwóch taśm odtwarzanych synchronicznie, z których jedna jest minimalnie spowolniona przez dociskanie palcem. Efekt ten jest zbiorem filtrów grzebieniowych, których charakterystyki są różne dla poszczególnych częstotliwości.

Taśma magnetyczna może zostać ponadto przesterowana, jeśli podane zostanie na nią zbyt duże pole magnetyczne. Przester taki często wykorzystywany jest w produkcji analogowej, ponieważ jego charakterystyczne brzmienie jest jedną z cech rozpoznawczych taśm magnetycznych.

Najprostszy procesor dźwięku, jaki można zbudować, jest zwykłym magnetofonem, którego prędkość odtwarzania można regulować. Otrzymuje się w ten sposób sygnał o zmienionym czasie trwania oraz częstotliwości. Nie można tego efektu wykorzystać „na żywo”, jednak można go nagrać i wykorzystać w czasie produkcji.

5. BADANIA SZUMU TAŚM MAGNETYCZNYCH

W celu sprawdzenia niedoskonałości taśm magnetycznych określony został stosunek sygnału do szumu nagrania na kasie magnetofonowej. Aby zbadać tę cechę, konieczne było zarejestrowanie sygnału odtwarzanego w formie cyfrowej, pozwalającej na łatwe przetwarzanie sygnału. W tym celu wyjście przenośnego odtwarzacza kaset zostało podłączone do wejścia liniowego karty dźwiękowej komputera. Nagrane zostały fragmenty utworów oraz ciszy występującej między piosenkami. Analiza przeprowadzona została w środowisku Matlab.

Do obliczenia stosunku sygnału do szumu wykorzystany został wzór:

$$SNR_{dB} = 10 \log_{10} \left(\frac{p_{S_RMS}}{p_{N_RMS}} \right)^2$$

gdzie:

- SNR_{dB} – stosunek sygnału do szumu,
- p_{S_RMS} – wartość skuteczna amplitudy sygnału,
- p_{N_RMS} – wartość skuteczna amplitudy szumu.

Zmierzone wartości zamieszczone zostały w tabeli 1.

Tabela 1

Zmierzone wartości poziomu stosunku sygnału do szumu

Próbka	p_{S_RMS}	p_{N_RMS}	SNR_{dB} [dB]
1	0,1162	0,0035	30,5
2	0,1838	0,0074	27,9
3	0,2369	0,0072	30,3
Średnia	–	–	29,6

Zmierzona średnia wartość SNR_{dB} wynosi 29,6 dB. Jest to bardzo mały odstęp sygnału od szumu. Szum jest wyraźnie słyszalny w nagraniu, jednak nie zniekształca on w sposób znaczący informacji przekazywanych za pomocą kasyety.

6. PODSUMOWANIE

Taśmy magnetyczne są ciągle żywym nośnikiem dźwięku, dały one początek dzisiejszej produkcji dźwiękowej. Mają zarówno wady, jak i zalety. Wiele programów DAW swój interfejs wzoruje na studyjnych urządzeniach taśmowych. Nośniki magnetyczne pozwalają w łatwy i precyzyjny sposób realizować produkcję dźwięku. Procesory dźwięku zbudowane na podstawie układów magnetycznych często były i są wykorzystywane przez wielu muzyków.

LITERATURA

- [1] Engel F.K. (ed.): Oberlin Smith and the Invention of Magnetic Sound Recording an Appreciation on the 150th Anniversary of the Inventor's Birth Documents, Publications, Letters, Biographical Sketches with an Introduction by Friedrich Karl Engel. Bensheim 2006, http://www.richardhess.com/tape/history/Engel--Oberlin_Smith_2006.pdf [data dostępu: 21.02.2016]

- [2] Valdemar Poulsen, https://web.archive.org/web/20091027123228/http://geocities.com/neveyaakov/electro_science/poulsen.html [data dostępu: 21.02.2016]
- [3] Method of recording and reproducing sounds or signals. US 661619 A, <https://www.google.com/patents/US661619> [data dostępu: 21.02.2016]
- [4] Daniel E.D., Mee C.D., Clark M.H.: Magnetic Recording: The First 100 Years, New York 1999
- [5] Urbański B.: Elektroakustyka, Kraków 1977
- [6] Urbański B.: Magnetyczny zapis dźwięków i obrazów, Warszawa 1964

STUDIO FOLEYA – IMITACJA DŹWIĘKU NA POTRZEBY PRODUKCJI ANIMACJI I FILMÓW

Foley studio – imitation of the sounds for the animation
and video production

Damian KOSZEWSKI, Karolina MARCINIUK

Politechnika Gdańska,
Wydział Elektroniki, Telekomunikacji i Informatyki,
damian.koszewski@gmail.com

STRESZCZENIE

Praca przybliży metody udźwiękawiania filmów, zasady projektowania, wyposażania oraz funkcjonowania studiów Foley'a na przykładzie animacji. Przedstawiono układ studia przystosowanego do nagrań dźwięków wraz z konfiguracją torów wizyjnych i fonicznych na przykładzie studia lektorskiego Katedry Systemów Multimedialnych Politechniki Gdańskiej. Omówiono etapy pracy z dźwiękiem w filmie/animacji. Pokazano przykładowe scenariusze użycia rekwizytów i stanowisk do imitacji dźwięków różnych powierzchni.

ABSTRACT

The aim of the paper was to present the methods of sound design for movie. The principles of design, equipment and operations on Foley studios were presented with sample animation. Configuration of the studio adapted for movie sound imitating along with audio-video configuration were presented based on the lector studio located at Multimedia Systems Department at Gdansk University of Technology. A few scenarios of sound design were shown.

1. WPROWADZENIE

Postprodukcja dźwięku do obrazu jest aktualnie niezbędnym etapem w produkcji filmów. Jednym z zabiegów jest dodawanie postsynchronów. Jest to ponowne nagrywanie dialogów, które nie zostały poprawnie zarejestrowane na planie. Wykonywane są synchronicznie z obrazem, co daje aktorowi możliwość odegrania sceny z odpowiednią manierą i uniknięcia efektu brzuchomówstwa (słyszymy głos aktora, ale jego usta się nie poruszają, efekt zauważalny często przy rozsynchronizowaniu się ścieżek wideo i audio). Drugi obszar postprodukcji obejmuje efekty dźwiękowe, które często są świadomie zaniedbywane podczas zdjęć filmowych.

Dźwięk w filmie i animacji można podzielić na trzy kategorie: narracja z dialogami, efekty dźwiękowe oraz muzyka [1]. Pierwsze dwie mają funkcję narracyjną, to znaczy składają się na nią dźwięki możliwe do zarejestrowania w obrazie. Trzecia występuje poza kadrem. Jej zadaniem jest wprowadzenie odpowiedniego klimatu sceny. Na potrzeby pracy skupiono się na drugiej kategorii ścieżki dźwiękowej w filmie.

Na efekt akustyczny związany z ruchem postaci składa się szereg elementów. Nie jest to jedynie odgłos kroków, ale również szeleszczenie ubrań, dźwięk ocierania się o ścianę, np. w sytuacji, gdy postać w filmie się skrada. Każdy z sygnałów źródłowych musi być dopasowany pod względem panoramy w obrazie (w przypadku dźwięków stereo i więcej) oraz poziomu natężenia.

Wprowadzenie efektów dźwiękowych do filmu czy animacji pozwala na przekazanie dodatkowych informacji dotyczących emocji bohatera, a tym samym wpływa na odbiorcę [2]. Udźwiękowanie filmów polega na dograniu ścieżki fonicznej składającej się z dźwięków charakterystycznych dla przedmiotów znajdujących się w danej scenie [3]. Odbywa się to w postprodukcji dźwięku, ponieważ w tym etapie pracuje się nad dokładną synchronizacją dźwięku i obrazu. Dźwięki te nazywane są efektami, ponieważ bardzo często nie naśladują one idealnie otoczenia, lecz je „przerysowują”, modyfikując dynamikę bądź kierując uwagę widza na ważny aspekt sceny [4].

2. UDŹWIĘKOWIANIE OBRAZU

2.1. STUDIA FOLEYA – BUDOWA

W przypadku dużych wytwórni filmowych studia Foley są zazwyczaj budowane od podstaw. W etapie projektowania i wdrażania uwzględnia się poszczególne stanowiska odtwórcze, różne powierzchnie do imitacji kroków, osobne przestrzenie do rejestracji odgłosów rekwizytów oraz przestrzenie magazynowe. Przydatną funkcjonalnością jest możliwość zmiany akustyki studia przez modyfikację właściwości akustycznych ścian, czyli zmianę materiałów oraz ich rozmiarów. W przypadku ograniczonego budżetu można wykorzystać obracane ustroje akustyczne, które z jednej strony mają powierzchnię odbijającą lub rozpraszającą dźwięk, z drugiej strony silnie pochłaniającą [4], co umożliwia dynamiczne zmiany czasu pogłosu w pomieszczeniu. Można je stosować w celu stałej zmiany akustyki do danej sceny lub symulacji płynnych zmian między scenami, wymagających zaangażowania dodatkowych osób.

Podkładanie dźwięku synchronicznie do obrazu wymaga stosowania sprzętu wizyjnego wnoszącego możliwie jak najmniej szumów. Typowo stosowane są projektory instalowane w pomieszczeniu reżyserskim rzucające obraz do studia. Ważną sprawą jest więc zapewnienie odpowiednich połączeń wizyjnych i fonicznych pomiędzy pomieszczeniem reżyserskim a studiami odtwórczymi. Wysokobudżetowe studia Foley korzystają z większej liczby pomieszczeń, w których nagrywa się dźwięki. Pozwala to na zróżnicowanie środowisk akustycznych w nagraniu. Wszystkie te zabiegi umożliwiają uzyskanie większej autentyczności wyprodukowanych dźwięków.

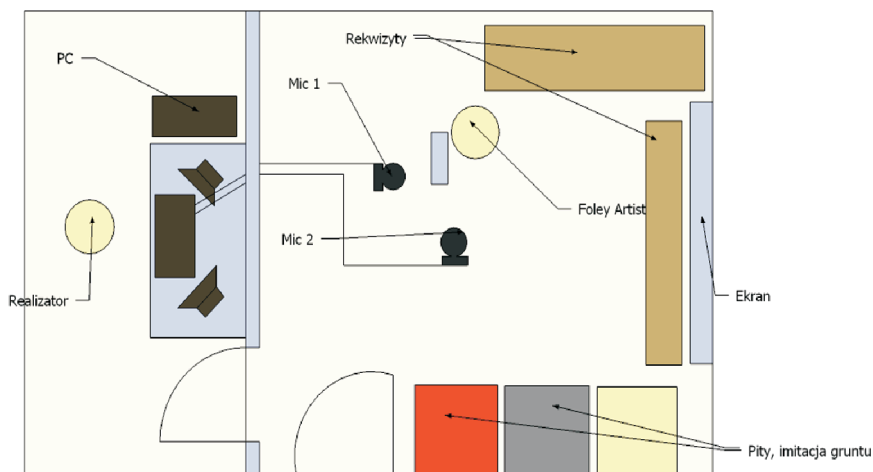
Do poprawnego funkcjonowania studia Foley wymagane jest zatrudnienie osób pracujących jako edytorzy dźwięku i artystów naśladowujących dźwięki. Kluczową rolę w każdym studiu odgrywa główny inżynier dźwięku. Jest on odpowiedzialny za funkcjonowanie całego obiektu oraz za przydzielanie zadań innym pracownikom.

Pomieszczenie reżyserskie w studiu Foley typowo wyposażane jest w stację roboczą wraz z odpowiednim oprogramowaniem DAW (*Digital Audio Workstation*). Przykładem takich zestawów może być komputer klasy Mac wraz z interfejsem M-Box firmy Avid oraz oprogramowaniem Pro Tools tej samej firmy. Wymagania obliczeniowe dla poszczególnych stacji roboczych zależą od stosowanych konfiguracji. Zaleca się procesory Intel i5, minimum 8 GB pamięci RAM, 15 GB przestrzeni dyskowej, połączenie sieciowe do instalacji oraz porty USB lub Firewire do komunikacji z interfejsami.

Przy nagrywaniu dźwięków warto stosować więcej niż jeden mikrofon. Często składowe dźwięki są inaczej rejestrowane przez mikrofony poszczególnych typów, które należy ustawić w różnej odległości od źródła. Końcowym efektem jest więc suma poszczególnych ścieżek, często po filtracji na poszczególnych pasmach. Dodatkową zaletą wykorzystywania dźwięków z różnych odległości od źródła jest możliwość późniejszego wykorzystania dźwięku w przypadku zmiany planu w filmie. Dobrze jest więc stosować miksery/konsole z przynajmniej ośmioma wejściami mikrofonowymi i zasilaniem Phantom (48 V).

2.2. STUDIO KATEDRY SYSTEMÓW MULTIMEDIALNYCH (KSM) POLITECHNIKI GDAŃSKIEJ

Studio KSM można podzielić na kilka stref (rys. 1). Jako studio Foleya wykorzystywane jest pomieszczenie studia lektorskiego. Podstawowymi elementami wyposażenia są pity zawierające różne materiały do symulacji podłoża. Pity to płytkie kwadratowe skrzynki wypełnione przykładowo kamieniami, liśćmi czy piaskiem (rys. 2).



Rys. 1. Schemat studia Foleya na przykładzie Katedry Systemów Multimedialnych Politechniki Gdańskiej



Rys. 2. Pity z różnymi materiałami do imitacji podłoża

Dodatkowo studio dysponuje płytami dla różnych typów posadzek, paneli i płytek ceramicznych o różnej podklejce. Pudełka z rekwizytami podzielone są według typów materiałów. Ze względu na podstawową funkcję pomieszczenia, elementy studia Foley są przenośne. Wymaga to więc zwrócenia uwagi na złożoną strukturę niektórych źródeł dźwięku. Odgłosy kroków inaczej zabrzmiały na luźno położonej płytce na podłodze, inaczej na płycie w całości przytwierdzonej do betonowej powierzchni. Dźwięki kroków na linoleum rozwiniętym na dywanie nie zabrzmiały tak samo jak w przypadku linoleum rozwiniętego bezpośrednio na podłożu. Często wymaga to kombinacji różnych materiałów.

2.3. BIBLIOTEKI DŹWIĘKOWE A NAGRYWANIE SAMODZIELNE

W celu uzupełnienia filmu lub animacji ścieżką dźwiękową z efektami można wykorzystać gotowe biblioteki próbek (darmowe lub płatne). Wynik jest jednak często bardzo niezadowolający i nie zawsze oddaje zamysł twórcy. Ponadto nawet profesjonalne zbiory bardzo rzadko charakteryzują się wystarczającą różnorodnością (np. dźwięki kroków). Przy wielokrotnym wykorzystaniu jednego pożądanego dźwięku powtarzalność nie może być zauważalna [5]. Większość dźwięków jest używana w filmie lub animacji wielokrotnie, więc naturalność dynamiki oraz brzmienia jest kluczowa.

Artyści studia Foley mają za zadanie opracować sposób naśladowania dowolnego zadanego brzmienia, nawet jeżeli jeszcze nie istnieje (np. odgłos miecza świetlnego z *Gwiezdnych wojen* [6]) i nagranie go wraz z późniejszą obróbką. Sposób naśladowania dźwięku często znacznie odbiega od tego, jak powstaje on w rzeczywistości (np. odgłos łamania ludzkich kości oddany za pomocą dźwięku towarzyszącego łamaniu warzyw). Niektóre z efektów zostają utrwalone w świadomości odbiorców kultury na tyle, że efekt filmowy staje się bardziej rozpoznawalny od rzeczywistego dźwięku [7] (np. wyjmowanie miecza z pochwy).

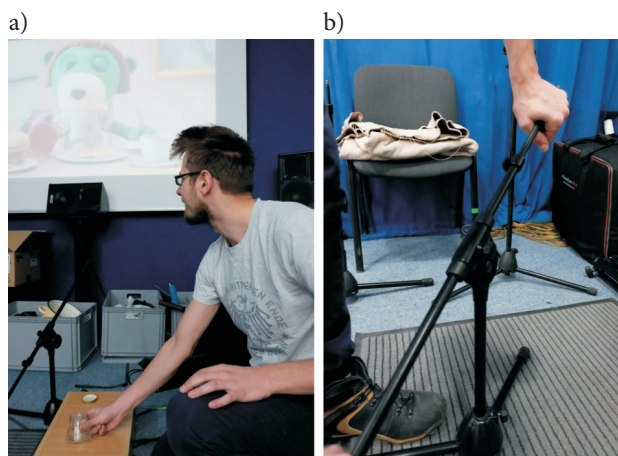
2.4. SCHEMAT NAGRAŃ

Udźwiękowanie filmu rozpoczyna się od obejrzenia go przez edytora nadzorującego wraz z reżyserem i głównym montażystą w celu określenia kolejności dodawania wymaganych dźwięków w każdej scenie. Efekty są rozpisywane na kilku ścieżkach, podawany jest ich czas i opis ujęcia [8–10]. Przed przekazaniem tego wykazu grupie artystów ścieżka dialogowa zostaje przeanalizowana i wyedytowana w celu wyodrębnienia dźwięków np. kroków, które udało się uchwycić przez mikrofony użyte przy nagrywaniu sceny, co często pozwala w znaczny sposób ograniczyć zakres zadań dla pracowników odpowiedzialnych za stworzenie właściwych efektów. Jeżeli jednak dźwięki tła mają pozostać oddzielone od dialogów, które np. zostaną wymienione na inną wersję językową, do dalszego wykorzystania mogą zostać użyte tylko fragmenty bez głosów aktorów [11]. Zdarza się również, że całe nagranie dźwiękowe gry aktorów zostaje usunięte w celu stworzenia go w studiu [6]. Po tym etapie artyści przystępują do pracy nad odtworzeniem pozostałych potrzebnych dźwięków, dopasowując swoją grę do konkretnych scen. Zadaniem edytorów jest zapis danych i odpowiednie zarządzanie układem ścieżek, a także wprowadzanie korekcy

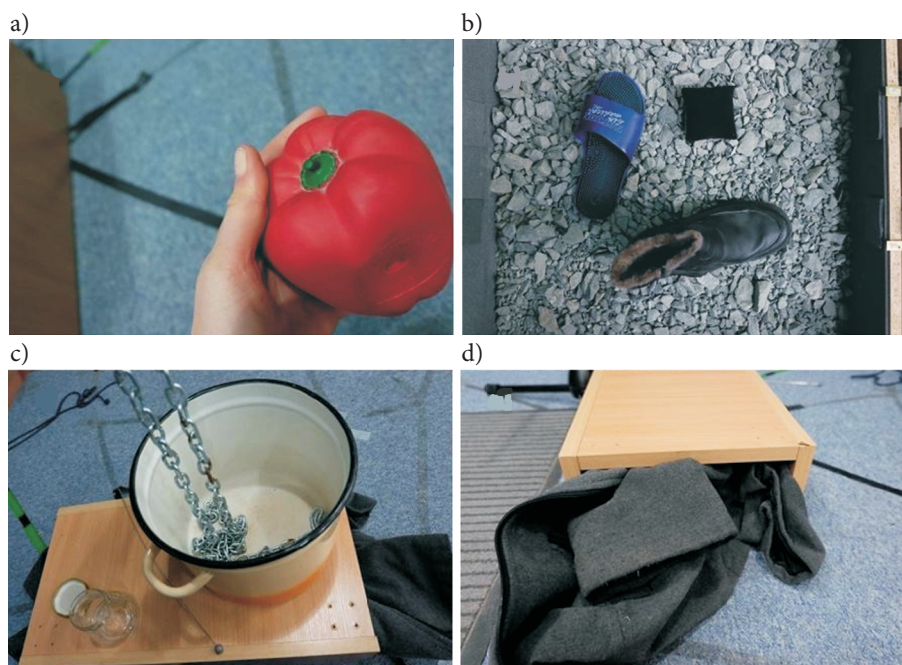
synchronizacji, kiedy jest to niezbędne. Kierują oni także pracą ludzi na scenie według opracowanego arkusza nagrań [8]. Wszystkie próby są nagrywane, najlepsze zostają przypisane do ścieżek podzielonych tematycznie, składają się one na podkład dźwiękowy do filmu [12]. Sceny powinny być nagrywane w całości do cięcia, aby dźwięki były spójne. Praca odbywa się zwykle w kilku etapach. Na samym początku dla każdej postaci nagrywane są odgłosy ruchów (ocieranie ubrań) [13], następnie ich kroki. Ostatnim etapem jest dodawanie pozostałych ścieżek uzupełniających film dźwiękami naśladującymi odgłosy wszystkich przedmiotów, niezbędnymi do dokończenia dzieła [14].

3. PRZYKŁADOWE SCHEMATY WYKORZYSTANIA

Do udźwiękowienia animacji dla dzieci użyto rekwizytów dostępnych w studio Katedry Systemów Multimedialnych Politechniki Gdańskiej (rys. 3, 4). Bohaterami bajki była rodzina małych zielonych stworków, a sama animacja została stworzona metodą poklatkową. Do imitacji odgłosów towarzyszących jedzeniu użyto półki pod komputer jako stołu oraz słoika i noża do naśladowania odgłosów naczyń. Nagrania przeprowadzono za pomocą mikrofonu AKG C4000 w odległości około 15–20 cm od półki. W scenie, w której Bohater A odkłada słoik na stół, uzyskano charakterystykę częstotliwościową z dużą ilością dolnych częstotliwości. Z powodzeniem udało się odfiltrować niepotrzebne częstotliwości za pomocą korektora graficznego (rys. 5). Podobny schemat był wykorzystany przy nagraniach innych dźwięków, dlatego zaleca się dobór odpowiedniego mikrofonu w zależności od nagrywanych dźwięków oraz akustyki pomieszczenia. Należy również rozważyć użycie filtra górnoprzepustowego (w mikrofonie) przy częstotliwości odcięcia 100 Hz, jednakże wówczas tracona jest bezpowrotnie informacja o dolnych częstotliwościach.



Rys. 3. Przykładowe schematy użycia rekwizytów: a) podkładanie synchroniczne dźwięku odgłosów jedzenia; b) dźwięk skrzypiącego statywu mikrofonowego jako odgłos otwierania drewnianych drzwi



Rys. 4. Przykładowe scenariusze użycia rekwizytów: a) ocieranie gumowej zabawki dla psów wydaje odgłos chodzenia po świeżym śniegu; b) wykorzystanie różnego obuwia pomaga przypisać bohaterom indywidualne dźwięki poruszania się; c) wrzucanie łańcucha do garnka symuluje odgłos pękających ziaren kukurydzy; d) wypełnienie szafki pod komputer różnymi materiałami tłumiącymi zmienia odgłosy uderzania w stół



Rys. 5. Charakterystyka częstotliwościowa: a) przed korekcją graficzną; b) po zastosowaniu korekcji graficznej

Podczas nagrań uzyskano ponad dwie godziny materiału, który następnie został poddany edycji i synchronizacji do obrazu. W tabeli 1 zostały umieszczone dźwięki, które nagrano do czwartego odcinka serii.

Tabela 1

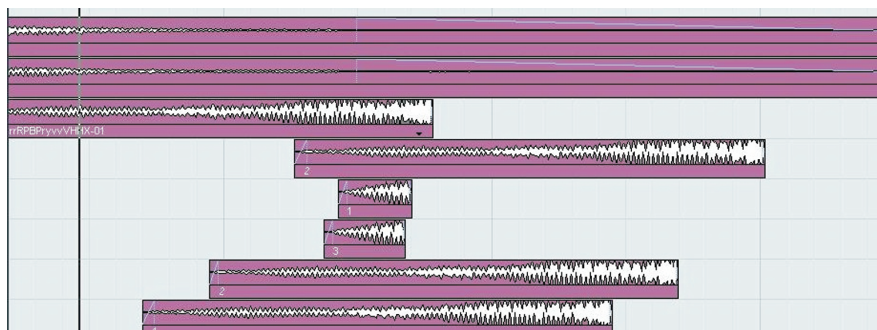
Lista dźwięków do odcinka 4

Kod czasowy	Efekt	Rzeczy użyte w scenie/działanie
1:07:550	odgłosy jedzenia bohaterów	uderzanie nożem o słoik, głośnie żucie
1:14:082	widelec po talerzu	uderzanie sztućcami o talerz
1:22:174	widelec uderza o słoik	uderzanie sztućcami o talerz
2:27:942	latająca mucha	pocieranie sznurka o oponę rowerową
2:39:610	świst powietrza, odganiające muchy	machanie płetwą do nurkowania
2:40:125	odgłosy materiałów – Bohater A	szeleszczenie materiałów
2:40:939	odgłosy materiałów – Bohater B	szeleszczenie materiałów
3:05:790	szamotanina, kroki	szeleszczenie materiałów, kroki
3:26:457	dźwięk wzbijania się Bohatera A w powietrze	machanie płetwą do nurkowania
3:27:070	Bohater B dmucha w wiatrak	potrząsanie kępką siana
3:29:695	Bohater C i Bohater A biegną	uderzanie butami o piasek
3:41:358	lot po pokoju Bohatera D	machanie płetwą do nurkowania
3:54:662	ładowanie na szafie	rzucenie butów zawiniętych w kurtkę
4:38:135	odgłosy krzesła i taboretu	uderzanie desek o siebie
4:48:138	skrzypienie drzwi	skrzypienie statywu mikrofonowego
5:05:044	skoki Bohatera E	uderzanie butami o piasek
5:13:500	uderzenie o podłogę trzech osób	rzucenie butów na piasek
5:23:017	szamotanina po podłodze	szeleszczenie materiałów, kroki
6:01:750	pukanie do drzwi	pukanie w deskę
6:29:348	dźwięki zastawy stołowej	uderzanie sztućcami o talerz
6:56:664	dźwięki obiadu	uderzanie sztućcami o talerz, głośnie żucie

4. CZĘSTE PROBLEMY PRZY PRACY W STUDIUM FOLEYA

Praca w studiu Foleya wymaga pomysłowości, zmysłu muzycznego i podstawowej wiedzy inżynierskiej z zakresu akustyki. Przygotowanie każdego dźwięku polega na wykorzystaniu gotowych sposobów na uzyskanie odgłosów lub eksperymentowanie w celu stworzenia ich od podstaw. Sztuka naśladowania jest dziedziną opartą na improwizacji. Wymaga od artysty inwencji i wyczucia, które często wypracowuje się przez wiele lat. W praktyce tylko niektóre dźwięki mogą być wygenerowane za pomocą przedmiotów widocznych w filmie. Bardzo często do imitacji dźwięku przedmiotu z obrazu wykorzystuje się przedmiot o zupełnie innym przeznaczeniu i wykonany z innego materiału. Czasem

jeden dźwięk jest wynikiem nakładania na siebie kilku osobno nagrywanych efektów dźwiękowych (rys. 6). Do tego celu akustycy stosują algorytmy korzystające z możliwości cyfrowego przetwarzania sygnałów w procesorach DSP (*Digital Signal Processing*). Przykładowo wykorzystywane są sztuczne pogłosy, kompresja, filtracja, korekcja oraz operacje spowalniania i przyspieszania czasu (*time stretching*) i zamiany częstotliwości podstawowej (*pitch shifting*).



Rys. 6. Wynik nakładania wielu warstw dźwięków

Kolejnym problemem jest synchronizacja nagranych dźwięków z obrazem. Bardziej doświadczonemu artyście wystarczy obejrzenie fragmentu filmu lub animacji, a następnie prawie bezbłędnie może odtworzyć pełną sekwencję efektów. W innym przypadku każde nagranie należy edytować i ręcznie przesuwac kolejne dźwięki na osi czasu, aby znalazły się na właściwym miejscu.

Ważną sprawą jest dobranie takiej liczby dźwięków, aby wspomagały odbiór obrazu, a nie go komplikowały. Nie zawsze wystarczy manipulacja poziomami głośności, odpowiednia selektywność jest częściej lepszym rozwiązaniem. Zbyt duża różnorodność efektów może odciągać zainteresowanie widza od kluczowego zdarzenia w filmie, zbyt mała ich liczba sprawi, że widz może pominąć zdarzenie mające wpływ na dalszą fabułę.

Ponadto artysta w studiu Foley'a powinien mieć podstawową wiedzę z dziedziny audio, szczególnie w kontekście parametrów mikrofonów i akustyki pomieszczeń. Wiedza w zakresie percepcji bodźców oraz ich oddziaływania na psychikę człowieka bywa niezastąpiona.

LITERATURA

- [1] Wołk-Łaniewski N., Napieralska J.: Analiza dźwięku w filmie, materiały z Uniwersytetu Muzycznego im. Fryderyka Chopina w Warszawie
- [2] Suchomski P.: Dźwięk w produktach multimedialnych, http://sound.eti.pg.gda.pl/student/ts/dzwiek_w_multimediach.pdf [data dostępu: 11.02.2016]
- [3] Sound, <http://www.mediaknowall.com/sound.html> [data dostępu: 11.02.2016]

- [4] Yewdall D.L.: *Practical Art of Motion Picture Sound*, Fourth Edition, Oxford, 2012
- [5] Skevos M.: What is Foley?, <http://www.sound-ideas.com/Foleymavart.html> [data dostępu: 11.02.2016]
- [6] Palermo T.: Zounds! What Sounds! Advice about Radio Sound Effects, Tony Palermo's RuyaSonic, http://www.ruyasonic.com/sfx_zounds.htm [data dostępu: 11.02.2016]
- [7] Foley Artistry: The ambient sound effects of Episode 1, Lucasfilm Ltd., <http://film-sound.org/Foley/episode1-Foley.htm> [data dostępu: 11.02.2016]
- [8] Borgia C.: How to Foley, School of Visual Arts, Department of Film, http://smcnetwork.org/system/files/smc2011_submission_111.pdf [data dostępu: 11.02.2016]
- [9] Przedpelska-Bieniek M.: *Dźwięk w filmie*, Warszawa 2006
- [10] Vilorio D.: You're what? Foley artist, *Occupational Outlook Quarterly*, 2011, vol. 55, 1, 34–35
- [11] Baker K.: What is Foley and Why Should You Care?, <http://angryfilmmaker.wordpress.com/2008/08/04/what-is-foley-and-why-should-you-care/> [data dostępu: 11.02.2016]
- [12] Ament V.: *The Foley Grail: The Art of Performing Sound for Film, Games, and Animation*, Oxford, 2009
- [13] Goodyer T.: Tim Goodyer investigates the overlooked subject of ADR and Foley in film and TV recording, <http://web.archive.org/web/20050306014850/http://audiomedia.com/archive/features/uk-0402/uk-0402-foley/uk-0402-foley.htm> [data dostępu: 11.02.2016]
- [14] Singer P.R.: Art of Foley, <http://www.marblehead.net/foley/> [data dostępu: 11.02.2016]