

**AKADEMIA GÓRNICZO-HUTNICZA  
IM. STANISŁAWA STASZICA W KRAKOWIE**



**AGH**

Wydział Informatyki, Elektroniki i Telekomunikacji  
Katedra Elektroniki

**PRACA DYPLOMOWA**

Inżynierska

**Krótkometrażowy film popularno-naukowy o technologiach  
mowy**

**Short educational film about speech technology**

Imię i nazwisko: Katarzyna Karkoszka

Kierunek studiów: Inżynieria akustyczna

Opiekun pracy: dr inż. Bartosz Ziółko

Kraków 2013

## **OŚWIADCZENIE AUTORA PACY**

OŚWIADCZAM, ŚWIADOMA ODPOWIEDZIALNOŚCI  
KARNEJ ZA POŚWIADCZENIE NIEPRAWDY, ŻE NINIEJSZĄ  
PRACĘ DYPLOMOWĄ WYKONAŁAM OSOBIŚCIE  
I SAMODZIELNIE, I NIE KORZYSTAŁAM ZE ŹRÓDEŁ  
INNYCH NIŻ WYMIENIONE W PRACY.

.....  
podpis

## Streszczenie

Celem niniejszej pracy było zaprezentowanie etapów realizacji projektu inżynierskiego, jakim był krótkometrażowy film popularno-naukowy o technologiach mowy. Produkcja przedstawiała projekty aktualnie realizowane przez Zespół Przetwarzania Sygnałów Akademii Górniczo-Hutniczej. W filmie zaprezentowano tworzone technologie przede wszystkim pod kątem ich zastosowania. Użyty język filmu pozwala na zapoznanie się z tą dziedziną nauki, szczególnie osobom nie posiadającym szczegółowej wiedzy technicznej w tym zakresie.

W kolejnych rozdziałach pracy opisane zostały trzy podstawowe etapy tworzenia projektu: pisanie scenariusza i realizacja zdjęć, montaż obrazu oraz realizacja i montaż dźwięku. W każdym przypadku najpierw przedstawiono zalecenia literatury przedmiotu co do konkretnych realizacji, a następnie przyjętą w trakcie tworzenia filmu metodę. Wybór każdej z możliwych opcji został uargumentowany i szczegółowo omówiony. W związku z kończącym kierunkiem studiów, największą uwagę autor skupił na rozdziale dotyczącym realizacji i montażu dźwięku. Skupiono się na takich procesach jak zgranie dźwięku na planie zdjęciowym, nagrania lektorskie oraz udźwiękowienie produkcji.

Wszystkie założone cele pracy zostały zrealizowane, a końcowy efekt, jakim jest krótkometrażowy film popularno-naukowy, jest dostępny na kanale TV AGH platformy *YouTube*, pod tytułem: *Technologie Mowy*.

# Spis treści

1. <u>WSTĘP</u> .....	5
2. <u>SCENARIUSZ I REALIZACJA NAGRAŃ</u> .....	6
2.1. PRZYGOTOWANIE SCENARIUSZA FILMU <i>TECHNOLOGIE MOWY</i> .....	6
2.2. TECHNICZNE ASPEKTY REALIZACJI ZDJĘĆ .....	8
2.3. ARTYSTYCZNE ASPEKTY REALIZACJI ZDJĘĆ .....	10
3. <u>MONTAŻ OBRAZU</u> .....	13
3.1. MONTAŻ-PODEJŚCIE PRAKTYCZNE .....	13
3.2. TECHNOLOGIE MONTAŻU FILMOWEGO .....	15
3.3. REALIZACJA PROCESU MONTAŻU W PROJEKCIE.....	17
4. <u>DŹWIEK W FILMIE</u> .....	20
4.1. TECHNIKI REALIZACJI DŹWIĘKU NA PLANIE .....	20
4.2. UDŹWIĘKOWIENIE I METODA FOLEY'A .....	22
4.3. REALIZACJA I MONTAŻ DŹWIĘKU W FILMIE <i>TECHNOLOGIE MOWY</i> .....	24
5. <u>ZAKOŃCZENIE</u> .....	31
<u>BIBLIOGRAFIA:</u> .....	32
<u>SPIS RYSUNKÓW:</u> .....	33

# 1. Wstęp

Głównym celem projektowej części pracy inżynierskiej było stworzenie krótkometrażowego filmu popularno-naukowego prezentującego projekty realizowane na Akademii Górniczo-Hutniczej, związane z technologiami mowy. Docelowo praca miała przedstawić potencjalnemu widzowi, czym zajmują się ta dziedzina nauki. Dzięki prezentacji przede wszystkim zastosowań technologii mowy w życiu codziennym, film miał być zrozumiały dla osób, które nie posiadają szczegółowej wiedzy technicznej z tego zakresu.

Opisowa część pracy miała starannie pokazać na czym opierały się działania podczas tworzenia filmu. Skupiono się na trzech podstawowych etapach, jakimi były: tworzenie scenariusza i realizacja zdjęć, montaż obrazu oraz realizacja i montaż dźwięku. W każdym przypadku najpierw przedstawiono zalecenia literatury przedmiotu co do konkretnych realizacji, a następnie przyjętą w trakcie tworzenia filmu metodę.

W rozdziale dotyczącym scenariusza i realizacji nagrań opisany został proces tworzenia właśnie scenariusza oraz szczegółowe metody jego realizacji podczas nagrań zdjęciowych. Zaprezentowano również na jakich technologiach tworzonych przez Zespół Przetwarzania Sygnałów AGH skupiono w filmie uwagę.

W następnym rozdziale przedstawiono współczesne metody montażu obrazu. Następnie szczegółowo uargumentowano i opisano wybraną metodę nieliniowego montażu elektronicznego. Opisano oprogramowanie użyte do tego celu oraz parametry, które były stosowane.

Ostatni z rozdziałów dotyczył najdłuższego etapu tworzenia filmu, jakim był montaż i realizacja dźwięku. Skupiono się tu zarówno na zgraniu dźwięku podczas planów zdjęciowych, jak i nagraniach lektorskich w studio oraz udźwiękowieniu ostatecznej wersji produkcji. Opisano wykorzystany sprzęt i oprogramowanie, wraz z uzasadnieniem użytych parametrów.

Końcowy efekt projektu inżynierskiego, czyli stworzony krótkometrażowy film popularno-naukowy o technologiach mowy, dostępny jest na kanale TV AGH platformy *YouTube*, pod tytułem: *Technologie Mowy*.

## 2. Scenariusz i realizacja nagrań

### 2.1. Przygotowanie scenariusza filmu *Technologie mowy*

Tworzenie scenariusza to pierwszy i najważniejszy, z punktu widzenia treści, element tworzonego filmu. Etapy jego tworzenia i elementy, które powinien zawierać ściśle zależą od gatunku filmowego, do którego będzie przeznaczony. Stworzony na potrzeby projektu film popularno-naukowy, ze względu na swój charakter raczej można zakwalifikować do filmów reportażowych czy dokumentalnych. Prezentuje on bowiem realne wydarzenia, więc nie miało tu miejsca klasyczne tworzenie postaci, bowiem one już istniały i należało jedynie odpowiednio je przedstawić.

Głównym celem projektu było stworzenie filmu popularno-naukowego, który będzie przedstawiał pracę Zespołu Przetwarzania Sygnałów AGH nad technologiami mowy. Produkcja miała docelowo prezentować te działania w taki sposób, aby były one zrozumiałe dla widzów, którzy nie wiedzą czym zajmują się ta dziedzina nauki. Autor postanowił więc, że opracowywane na Akademii Górniczo-Hutniczej techniki będą przedstawiane poprzez ich zastosowanie. W ten sposób widz będzie mógł skojarzyć, że spotyka się z wykorzystaniem tych technologii na co dzień, tylko nie wie jak one powstają.

Po konsultacjach postanowiono przedstawić 4 technologie, nad którymi pracuje Zespół Przetwarzania Sygnałów. Należą do nich:

- anotator, czyli program stworzony na potrzeby rozpoznawania mowy,
- identyfikacja mówców,
- automatyczne protokołowanie,
- systemy wielomikrofonowe.

Pierwsza z prezentowanych technologii dotyczy stworzonego na potrzeby systemów rozpoznawania mowy, programu zwanego anotatorem. Podstawowym celem programu jest dzielenie dłuższych fragmentów mowy na mniejsze, które są następnie poddawane automatycznemu rozpoznawaniu mowy. Pozwala to na dokładne określenie czasu nagrania, w jakim znajdują się odpowiednie słowa. Program pomaga w późniejszym trenowaniu różnego rodzaju systemów do rozpoznawania mowy i identyfikacji mówców.

Kolejne zaprezentowane w filmie zastosowanie technologii mowy, to system pozwalający na identyfikację mówców. Opiera on swoje działanie na indywidualnych cechach każdego głosu, takich jak barwa, ton czy własności biometryczne. System posiada dużą skuteczność i ma duży potencjał związany z wykorzystaniem komercyjnym. Tego typu programy mogą bowiem zastąpić używane współcześnie systemy bezpieczeństwa służące do identyfikacji klientów banków, lub posłużyć do działań śledczych policji, gdy głos danego przestępcy znajduje się w bazie.

Następnym przedstawionym projektem tworzonym dla mowy polskiej jest system automatycznego protokołowania. Jego głównym celem jest zamiana wyrazów wypowiedzianych przez osobę mówiącą w tekst. Naturalnym sposobem porozumiewania się ludzi jest mowa, natomiast komputer nieporównywalnie lepiej radzi sobie z komunikacją za pomocą symboli. I właśnie w celu stworzenia jak najdogodniejszego systemu komunikacji dla obu stron, tworzone są tego typu programy. Automatyczne protokołowanie może zastąpić klasyczną klawiaturę, bądź też służyć do wydawania poleceń różnym urządzeniom.

Ostatnie prezentowanie w filmie zastosowanie technologii mowy tyczy się systemów wielomikrofonowych. Systemy te służą do lokalizacji źródła dźwięku i opierają swoje działanie na różnicach fazowych występujących w rejestrowanym sygnale. Z tą samą metodą działania mamy do czynienia podczas nagrań stereofonicznych tyle tylko, że w oparciu głównie o dwa mikrofony. Tutaj mamy do czynienia nawet z systemami wykorzystującymi 16 mikrofonów, co pozwala na bardzo precyzyjną lokalizację źródła. Urządzenia takie mogą być stosowane m.in. w systemach bezpieczeństwa, czy podczas różnego rodzaju konferencji, gdzie nie ma możliwości podpięcia osobnego mikrofonu dla pojedynczego mówcy.

Każda z prezentowanych technologii stanowiła osobną scenę w filmie, składającą się z wypowiedzi członka Zespołu Przetwarzania Sygnałów odpowiedzialnego za pacę nad danym systemem. Pokazano również urządzenia niezbędne do pracy przy każdym z systemów. W filmie zawarto również sondę uliczną, w której pytano przechodniów (głównie studentów AGH), czym zajmują się technologie mowy. Celem tego było zaprezentowanie poziomu wiedzy przypadkowych osób na temat tej konkretnej dziedziny, co miało stanowić odpowiedni wstęp dla dalszych części produkcji.

## 2.2. Techniczne aspekty realizacji zdjęć

Wspominając o technicznych aspektach realizacji obrazu należy wspomnieć o początkach jego zapisu. Tutaj najważniejszym wydarzeniem, które wpłynęło na rozwój kinematografii, było pokazanie światu w 1895 roku przez braci Lumiere, kinematografu. Bracia są często mylnie przedstawiani jako pierwsi wynalazcy tego urządzenia. Podczas gdy, oni jedynie znacznie ulepszyli jego działanie i przedstawili możliwości urządzenia większej publiczności [14]. Kinematograf było to urządzenie, które pozwalało na realizację i ekranową projekcję ruchomych obrazów. Nie da się ukryć, że zrewolucjonizowało ono podejście do robienia zdjęć, które teraz mogły jeszcze bardziej oddawać rzeczywistość. Bardzo istotne było również wynalezienie światłoczułego nośnika jakim była taśma filmowa, która pozwalała na zapis filmu w postaci pojedynczych klatek (kadru) odtwarzanego później przez projektor [15].

Współcześnie mamy do czynienia głównie z cyfrową wersją obrazu. Bardzo rzadko stosuje się kamery, które jako nośnika używają jeszcze taśmy filmowej. Nie trwa to jednak długo, zjawisko to raczej zaniknęło dopiero kilka lat temu, chociaż już kilkadziesiąt lat wcześniej znany był proces digitalizacji obrazu. Wiąże się to ze znacznym rozwojem techniki komputerowej, dzięki której współczesne komputery są szybkie i posiadają duże możliwości pamięciowe. Proces digitalizacji obrazu w teorii nie różni się znacząco, od tego przeprowadzanego na sygnałach jednowymiarowych. Obraz cyfrowy jest uzyskiwany poprzez próbkowanie prostokątnej siatki punktów obrazu analogowego, z odpowiednią rozdzielczością. Każdy z punktów siatki, cechuje się daną wartością tonalną (reprezentowaną za pomocą ciągu binarnego), która może być biała, czarna, posiadać odcień szarości lub kolor [15].

Do realizacji zdjęć podczas projektu używano cyfrowej kamery firmy Sony, model NEX-FS100. Posiada ona matrycę Super 35mm, o rozdzielczości 3.37 miliona pikseli, która zapewnia dużą plastyczność obrazu. Do kodowania zastosowaną w niej kodeka AVCHD, który jest formatem do zapisu sygnału o dużej rozdzielczości, kamera bowiem nagrywa w standardzie HD [16]. Nie da się ukryć, że urządzenie to w rękach doświadczonego operatora pozwala na tworzenie w pełni profesjonalnych obrazów.



Należy pamiętać że kamera to nic innego jak aparat fotograficzny robiący zdjęcia z określoną częstotliwością, najczęściej 24 klatki/sekundę. W formach telewizyjnych wykorzystuje się czasem także 50klatek/sekundę. Najważniejszym elementem aparatu jest więc obiektyw, którego odpowiedni dobór i ustawienia wpływają na głębie obrazu, czy wyrazistość prezentowanych elementów. Do wyboru mamy 4 podstawowe rodzaje obiektywów [Kingdon 2007, s. 174-176]:

- obiektywy szerokokątne – pozwalają na kręcenie obrazów na szerokim polu (np. tzw. rybie oko),
- obiektywy długoogniskowe – sprawiają, że obraz jest ciasny, jakby zbity, ostry na pierwszym planie,
- obiektywy średnioogniskowe – są niejako środkiem pomiędzy obiektywem szerokokątnym i długoogniskowym, pokazują ostrość obiektów w taki sam sposób, niczego nie wyróżniają,
- obiektywy zmiennoogniskowe – posiadają grupę soczewek, które poruszają się wzdłuż osi obiektywu, umożliwiając płynną regulację długości ogniskowej.

Podczas realizacji zdjęć do filmu projektowego wykorzystano wszystkie rodzaje obiektywów. Szerokokątne idealnie sprawdziły się do kręcenia ujęć z dużymi obiektami, jak budynek AGH, czy sala pełna studentów w sali komputerowej. Pozwoliło to na wyeksponowanie powierzchni. Najczęściej używanym rodzajem obiektywu był obiektyw średnioogniskowy, który był stosowany podczas wszystkich wywiadów. Odwzorowuje on rzeczywistość w sposób najbardziej naturalny, dlatego znalazł zastosowanie właśnie w tych ujęciach. Wszystkie detale, czyli zbliżenia obiektów i urządzeń, które stanowiły ujęcia do przebitek były kręcone obiektywem długoogniskowym. Daje on bowiem możliwość uwydatnienia konkretnego obiektu poprzez pozbawienie ostrości reszty ujęcia. Ostatni z rodzajów obiektywów, czyli ten o zmiennej ogniskowej był używany podczas sondy ulicznej zamieszczonej na początku filmu. W tej sytuacji nie można było przewidzieć w jakiej odległości od obiektywu znajdzie się ankietowany, a liczył się czas, dlatego z góry podjęto decyzję o zastosowaniu tego obiektywu. Inne parametry pracy kamery jak balans bieli i ostrość były dopasowywane do odpowiednich sytuacji, w szczególności naświetlenia pomieszczenia i warunków atmosferycznych.

### 2.3. Artystyczne aspekty realizacji zdjęć

Każdy film, niezależnie od rodzaju składa się ze scen, a sceny z ujęć. Ujęcie jest podstawowym elementem rozgraniczonym podczas realizacji zdjęć włączeniem i wyłączeniem kamery, a podczas montażu kolejnymi cięciami. Scenę natomiast określa miejsce, w której została nakręcona. Nowe miejsce równa się nowa scena [Mascelli 2008, s. 22-23]. Jeżeli kadr jest statyczny, czyli operator nie porusza kamerą, w danym ujęciu jego plan jest niezmienny. Wielkość planu wraz z kątem filmowania obiektów i wysokością kamery stanowią kąt widzenia kamery, który ściśle wpływa na odbiór filmu przez widza [Mascelli 2008, s. 40].

W sumie wyróżnia się około 9 rodzajów planów. W pracy jednak zostaną opisane tylko te, który były wykorzystywane w filmie. Najczęściej zobaczyć w nim można półzbliżenie, które było wykorzystywane podczas wszystkich wywiadów. Poniżej na rysunku przedstawiono tego typu ujęcie.



**Rys. 1. Półzbliżenie**

*Źródło: opracowanie własne*

Pokazuje ono postać od klatki piersiowej do czubka głowy, z małą przestrzenią powietrza nad głową. Dość często można zauważyć również plan zwany detalem, który był używany do pokazania urządzeń, komputerów, czy innych istotnych fragmentów jakiejś czynności. Plan średni, czyli od pasa w górę postaci również był wykorzystywany do ujęć służących jako

przebitki. Ma on tę zaletę, że w kadrze mieści się kilka postaci, których gesty i wyrazy twarzy są wyraźne dla widza. Na początku filmu można także zauważyć przedstawiony budynek AGH, który był kręcony w planie ogólnym, aby pokazać niejako miejsce akcji.

Niezwykle istotnym aspektem podczas filmowania jest kąt pod jakim znajdują się elementy w kadrze. Aby widoczna dla widza była przynajmniej pozorna trójwymiarowa rzeczywistość na dwuwymiarowym ekranie, należy pokazać przynajmniej dwie płaszczyzny danej bryły. W ten sposób uzyskuje się głębie, którą można również otrzymać za pomocą prawidłowego oświetlenia, ruchu kamery, czy odpowiedniego obiektywu. Jednak filmowanie obiektów pod odpowiednim kątem jest metodą najprostszą i najskuteczniejszą [Mascelli 2008, s. 54-55].

Ostatnim elementem składającym się na kąt widzenia kamery jest jej wysokość. Wpływa on na dramaturgie oraz artystyczne aspekty prezentowanej historii. Wyróżnia się kilka podstawowych ułożeń i perspektyw, które wpływają na odbiór przez widza. Należą do nich [Mascelli 2008, s. 56-78]:

- kamera na wysokości oczu,
- perspektywa ptasia,
- perspektywa żabia,
- kamera przekrzywiona.

Pierwsze z wymienionych ułożeń, czyli kamera na wysokości oczu postaci sprawia wrażenie, że postać utrzymuje z widzem kontakt wzrokowy. Można też wprowadzić wariant, w którym kamera jest ustawiona na wysokości oczu obserwatora o przeciętnym wzroście. Tego typu zabieg stosuje się często w filmach edukacyjnych, aby widz mógł poczuć się jakby był elementem akcji. Tego typu perspektywa był również wykorzystywana w filmie projektowym, podczas kręcenia wszystkich wywiadów. Dwie kolejne perspektywy, czyli ptasia i żabia, są dość znane i często używane w fotografii. Ich głównym zamierzeniem jest uzyskanie właśnie efektu artystycznego. Pokazanie obiektu z niecodziennego punktu widzenia. Bardzo ciekawy efekt można uzyskać stosując ostatnią z wymienionych perspektyw, czyli przekrzywioną kamerę. Zazwyczaj polega to na przesunięciu osi pionowej kamery w stosunku do osi pionowej obiektów, co sprawia że linie w kadrze stają się ukośne. Stosowane są zazwyczaj w scenach niestandardowych, pokazujących na przykład rozchwianie psychiczne bohatera [Mascelli 2008, s.56-77].

Podczas realizacji zdjęć należy pamiętać zarówno o tych artystycznych, jak i technicznych aspektach. Bowiem tylko odpowiednie połączenie wiedzy i doświadczenia z obu tych obszarów pozwolą na tworzenie pięknych ujęć, które jednocześnie będą prezentowały to co w filmach najważniejsze, czyli odpowiedni przekaz.

## 3. Montaż obrazu

### 3.1. Montaż-podejście praktyczne

Po realizacji i zgraniu zdjęć film stanowi jedynie luźną kolekcję ujęć. Zgodnie z definicją M. Hendrykowskiego *montażem nazywamy proces wyboru kombinacji elementów wizualnych i dźwiękowych komunikatu kinematograficznego w spójną całość* [Hendrykowski 1999, s. 103]. Jest to więc proces, który wybiera, przycina i dopasowuje ujęcia w taki sposób, aby pasowały do opowiadanej historii. Zadaniem montażysty jest stworzenie jak najlepszej opowieści z dostępnych materiałów i nadanie jej odpowiedniego tempa [Mascelli 2008, s. 233].

Z punktu widzenia realizacji nagrań na planie można wyróżnić montaż chronologiczny i kompilacyjny. Ten pierwszy cechuje się tym, że scena złożona z takich ujęć sprawia wrażenie rozgrywającej się bardzo płynnie i nie eksponuje nadmiernie cięć. Tego typu montaż jest głównie wykorzystywany w produkcjach filmowych. Ten drugi natomiast jest ze względu na formę stosowany w różnego rodzaju produkcjach telewizyjnych, czy dokumentalnych. W montażu kompilacyjnym wykorzystuje się bowiem ujęcia z różnych źródeł, również te pochodzące z materiałów archiwalnych. Ich kompatybilność jest zachowana zazwyczaj tylko dzięki lektorowi, opowiadającemu co się dzieje w filmie [Mascelli 2008, s. 235]. Tego typu montaż został wykorzystany również w realizowanym na potrzeby pracy filmie. We wstępie przedstawiane są bowiem zdjęcia pochodzące z różnych źródeł, które przedstawiają historię prac nad technologiami mowy.

Kolejnym istotnym elementem procesu montażu, po wyborze jego rodzaju, jest realizacja cięć. Decyzja o tym, w którym momencie powinno nastąpić cięcie jest bardzo istotna z punktu widzenia celu produkcji. Operacja ta bowiem decyduje o tempie filmu, zaangażowaniu widza oraz ogólnie o jakości prezentowanego obrazu. Zdaniem W. Murch'a cięcie powinno spełniać sześć podstawowych kryteriów [Murch 2006, s. 36-37]:

- odzwierciedlenie nastroju chwili,
- popychanie historii do przodu,
- pojawienie się w momencie, który jest ciekawy pod względem rytmiki,
- zapewnienie obserwowania akcji przez widza,
- szanowanie zasady „planarności”, czyli dwuwymiarowości płaszczyzny ekranu,
- szanowanie zasady ciągłości trójwymiarowej rzeczywistej przestrzeni.

O zasadach tych należy pamiętać zarówno podczas montażu, ale także na wcześniejszych etapach tworzenia filmu. Kierowanie się nimi pozwoli niejednokrotnie na skrócenie czasu podejmowania decyzji montażowych. Istotna jest również ich kolejność, która oznacza stopień ważności. W razie konieczności zrezygnowania z którejś z nich, zawsze należy rezygnować z tej o mniejszej wadze dla realizowanego projektu. Najważniejsze są emocje, opowiadana historia i rytm.

Jeszcze jednym istotnym procesem podczas montażu filmowego jest łączenie ze sobą nie tyle poszczególnych ujęć, co już całych scen. Służy to głównie skracaniu czasu i przestrzeni w filmie, w taki sposób, aby widz nie był tym faktem zbyt zaabsorbowany, aby przebiegało to dla niego w sposób naturalny. Najprostszym sposobem, który jest często wykorzystywany w filmach dokumentalnych, czy popularno-naukowych jest łączenie scen za pomocą napisów. Widz zapoznaje się z tym, co następuje w kolejnej scenie dzięki napisowi umieszczonemu w kadrze. W ten sposób wprowadzany jest niejako do następnej części. W łączeniu scen za pomocą obrazu wykorzystywane są także różnego rodzaju ściemnienia, przenikania, lub tak zwane roletki, czyli w najprostszym wydaniu linia, która przesuwając się po ekranie zastępuje jedno ujęcie drugim [Mascelli 2008, s. 217-219]. Wszystkie tego typu narzędzia dostępne są każdym programie do obróbki plików wideo, a ich wybór ściśle zależy od realizowanego filmu. Dane przejście może w jednym gatunku wyglądać profesjonalnie, a w drugim groteskowo. Współcześnie, jeżeli jest to tylko możliwe, najczęściej wykorzystuje się bardziej błyskotliwe łączenie m.in. za pomocą odpowiedniego ujęcia, najczęściej zbliżenia na dany detal występujący w kadrze, czy za pomocą ruchu kamery.

Drugim często wykorzystywanym sposobem montowania scen jest łączenie za pomocą dźwięku. Do tego celu można wykorzystać głos lektora dochodzący z *offu*, monolog aktora lub fragment muzyki, który zaczyna się w poprzedniej scenie i jest kontynuowany w kolejnej. Istnieje wiele pomysłów dźwiękowych, które pomogą w dalszym poprowadzeniu akcji. Wszystko to służy wyeliminowaniu pustych miejsc w filmie, czy nagłych przejść, których nie da się uniknąć ze względu na charakter obrazu. *Voice-over*, bo tak fachowo nazywa się głos lektora mówiącego zza kadru, najczęściej wykorzystywany jest w reportażach, dokumentach, wiadomościach telewizyjnych oraz filmach, które wykorzystują materiały archiwalne [Mascelli 2008, s. 225-227]. W realizowanym na potrzeby pracy projekcie został wykorzystany zarówno sposób łączenia napisami, jak i dźwiękiem. Więcej na temat samej realizacji tego procesu przedstawione zostanie w dalszej części pracy.

## 3.2. Technologie montażu filmowego

Technologie montażu to inaczej techniczne sposoby wykorzystywane do łączenia ze sobą dwóch odrębnych ujęć. Na przestrzeni ponad wieku miało miejsce kilka znaczących rewolucji w tej dziedzinie, zazwyczaj służących usprawnieniu tego procesu.

Pod koniec XIX i na początku XX wieku podstawowymi narzędziami używanymi podczas montażu filmowego była przewijarka, para nożyczek i szkło powiększające. Była to era montażu manualnego, gdzie czas był liczony za pomocą odległości między palcami montażystki, a taśma filmowa składana była przy użyciu spinaczy do papieru. Montażownia przypominała wtedy trochę zakład krawiecki, gdzie praca polegała na ciągłych przymiarkach i poprawkach, tyle tylko że materiałem była taśma filmowa [Murch 2006, s. 93]. Nie było możliwości jednoczesnego edytowania i oglądania danego materiału.

Pierwszy przełom w technologii montażu filmowego nastąpił w roku 1924. Wtedy to wynaleziono w Stanach Zjednoczonych Maviolę, maszynę do mechanicznego montażu filmu. Urządzenie było pierwotnie projektowane jako domowy projektor, jednak ze względu na wysoką cenę nie cieszyło się ono zbyt dużą popularnością. Po wprowadzeniu kilku znaczących poprawek zaczęto wykorzystywać je jako maszynę do edycji taśmy filmowej, która pozwalała na bieżąco oglądać zmiany dokonywane na stole montażowym [9]. Na początku swojego istnienia urządzenie nie miało wielu zwolenników. Uważano je za zbyt kosztowne i hałaśliwe, a jego podstawową zaletę, jaką była możliwość przeglądania obrazów w ruchu, ignorowano. Momentem, w którym ta powszechna opinia uległa zmianie było wynalezienie w 1927 roku techniki zapisu dźwięku. Dwugłowicowa Moviola odtwarzała bowiem jednocześnie obraz i dźwięk. Urządzenie posiadało swoje europejskie odpowiedniki, później już bardziej zaawansowane, jak niemiecki Steenback, włoski Pevostem, czy francuski Moriton [Murch 2006, s. 94]. Montaż mechaniczny do końca XX wieku gościł we wszystkich studiach postprodukcyjnych na świecie i pomimo rewolucji komputerowej, wielu montażystom ciężko było rozstać się z zasłużoną maszyną.

Montaż mechaniczny był coraz częściej zastępowany montażem elektronicznym dopiero na początku XXI wieku. Skąd to opóźnienie, skoro pierwsze poważne próby wykorzystywania metody posługującej się oprogramowaniem komputerowym miały miejsce już w latach 80-tych? Przede wszystkim wczesne nieliniowe systemy elektroniczne, a raczej mechaniczno-elektroniczne, bowiem na początku były wykorzystywane jako urządzenia sprawujące kontrolę nad bankami maszyn analogowych, miały znaczące wady. Najważniejsza z nich to bardzo ograniczona pamięć, co w przypadku ogromnej ilości materiału wideo było

dość znaczącym problemem. Zastrzeżenia były również do ówczesnej jakości obrazu i wysokich kosztów wprowadzania do systemu materiałów filmowych [Murch 2006, s. 107-111]. Nadal bowiem do zapisu obrazu w kamerach używano taśmy filmowej, której zawartość była wprowadzana do komputera za pomocą procesu digitalizacji. Sytuacja zmieniała się wraz z rozwojem komputerów, ich możliwości obliczeniowych i pojemnościowych. Wtedy to zaczęto dostrzegać podstawowe zalety tego typu systemów i wykorzystywano go już do montażu czysto elektronicznego. Nie oznacza to wyparcia z rynku taśmy filmowej, która w profesjonalnych studiach filmowych wykorzystywana jest do dziś. Cały proces tworzenia filmu do momentu jego dystrybucji kinowej, odbywa się w technologii cyfrowej, jedynie po ostatecznym montażu nośnikiem znowu staje się taśma filmowa. Inaczej jest w przypadku branży telewizyjnej, gdzie już dawno zrezygnowano z tego nośnika.

Do podstawowych zalet montażu elektronicznego można zaliczyć [Murch 2006, s. 99-103]:

- większą prędkość, związaną przede wszystkim ze swobodnym dostępem do materiału zgromadzonego na dysku twardym i łatwością wyszukiwania poszczególnych ujęć, co w przypadku montażu mechanicznego było bardzo czasochłonnym zadaniem i wymagało zatrudnienia większej ilości osób do pracy w montażowni,
- możliwość przeglądania różnych wersji materiału, która podczas montażu mechanicznego była praktycznie niemożliwa, bowiem raz pocięta taśma nie dawała się łatwo złożyć do stanu pierwotnego, wymagane było kosztowne wykonywanie kopii zapasowych, system elektroniczny zapamiętuje zaś próbę montowania każdej nowej wersji sceny,
- zaawansowane wykorzystanie dźwięku, czyli możliwość operacji na wielu ścieżkach dźwiękowych jednocześnie i brak konieczności ich ciągłej synchronizacji z obrazem, ponieważ system automatycznie utrzymywał dźwięk w odpowiedniej relacji do obrazu, niezależnie od ilości wprowadzanych zmian,
- integracja z elektronicznymi efektami specjalnymi, które można od razu zobaczyć na montowanym materiale.



Współcześnie wszystkie te zalety i jeszcze wiele innych decydują o popularności montażu elektronicznego, zarówno w profesjonalnych, jak i amatorskich produkcjach. Biorąc pod uwagę również malejące ceny tego typu programów, coraz więcej osób, przy odrobinie talentu i doświadczenia, może pozwolić sobie na tworzenie domowych produkcji na profesjonalnym poziomie.

Na początku ery nieliniowego montażu elektronicznego swój prym wiodły takie oprogramowania jak Avid, Lightworks, EditDroid, czy CMX. Współcześnie do najpopularniejszych programów do elektronicznego montażu filmowego zaliczyć można Final Cut Pro, dedykowany do systemów operacyjnych firmy Apple, Lightworks, Adobe Premiere i Sony Vegas działający pod systemem Microsoft. Niezbędne wymagania systemowe do obsługi tych programów są aktualnie już tak powszechne, że nawet najniższej klasy notebook poradzi sobie z ich działaniem. Jediną przeszkodą może być czas w jakim program, w zależności od komputera, będzie przetwarzał informacje. Wszystkie te systemy oferują możliwość pracy na teoretycznie nieograniczonej liczbie ścieżek wideo oraz ścieżek dźwiękowych, posiadają zestawy profesjonalnych filtrów do edycji koloru, możliwość dodawania efektów w czasie rzeczywistym oraz współpracę z dostępnymi formatami i kodekami, a także obsługują zewnętrzne wtyczki z dodatkami.

W następnym podrozdziale szczegółowo opisany, pod kątem jego użytkowania, zostanie program Sony Vegas. To on służył bowiem jako podstawowe narzędzie montażowe podczas realizacji filmu.

### **3.3. Realizacja procesu montażu w projekcie**

Praca nad montażem obrazu dla realizowanego filmu, była jednym z dłuższych procesów, nie licząc realizacji i montażu dźwięku. Podczas tego etapu należało pamiętać, że mamy do czynienia z filmem reportażowym, który rządzi się nieco innymi prawami i zasadami montażowymi, niż film fabularny. W związku z tym, że w przypadku niniejszego projektu cały scenariusz opiera się na wypowiedziach osób występujących w filmie, najważniejsze było, aby owe wypowiedzi były składne i merytorycznie sensowne.

Programem używanym do obróbki plików wideo był Sony Vegas pro 11, który jest profesjonalnym narzędziem do cyfrowej edycji tych plików. Daje on możliwość wykorzystywania szerokiej gamy efektów związanych z korekcją barwy, animacjami, napisami, czy przejściami między ujęciami. Poniżej zamieszczono wygląd głównego okna, w którym pracowano podczas montażu.



**Rys. 2. Okno pracy programu Sony Vegas**

*Źródło: opracowanie własne*

Na powyższym rysunku można zauważyć podstawowe części z jakich składa się prawie każdy tego typu program. W prawym górnym rogu znajduje się okno wideo, które pokazuje film w sposób ciągły z prędkością zadaną przez operatora podczas kręcenia (w tym przypadku 24klatki/sekundę). Na dole widoczne są ścieżki wideo oraz audio, na których dokonuje się wszystkich operacji cięcia, łączenia, przejść itd. Na górze po lewej stronie jest natomiast widoczne okno efektów, w którym na przedstawionym rysunku znajdują się korektor barwy.

Praca nad montażem została podzielona na kilka podstawowych etapów. Pierwszy etap pracy to podział materiału filmowego na sceny i jego selekcja. Zgodnie ze stworzonym wcześniej scenariuszem, po zgraniu całego materiału z kamery, został on podzielony na 8 podstawowych scen. Następnie wszystkie ujęcia zostały obejrzone zarówno pod kątem jakości obrazu, jak i zgodności z treściami zawartymi w scenariuszu. W przypadku ujęć, które posiadały kilka dubli dokonano wyboru najlepszych wersji. Taka pierwsza faza pracy jest ważna ze względu na organizację procesu montażu i uniknięcie ewentualnych błędów, np. takich jak złe dopasowanie przebitki do danego materiału, czy brak odpowiedniej ilości materiału.

Kolejny etap to montaż poszczególnych setek, czyli nagrań zawierających ostateczną wersję obrazu i dźwięku. Biorąc pod uwagę formę produkcji, jaką niewątpliwie jest film reportażowy, najważniejsze na początku jest odpowiednie złożenie kolejnych wypowiedzi, scena po scenie. Na tym etapie zostały wycięte wszystkie niedociągnięcia nagranych dźwięków jak na przykład zająknięcia wypowiadających się osób. Cały czas trzeba było również brać pod uwagę jakość obrazu, jednak w związku z tym, że ujęcia wywiadów były statyczne, nie było z nimi większego problemu. Dokonano więc wyboru ujęć wszystkich 6 wywiadów, a następnie połączono każdą scenę w całość. Ważne było zachowanie proporcji szczególnie między setkami a wypowiedziami lektora z *offu*, który wprowadzał do każdej nowej sceny. Zbyt częsta zmiana głosów byłaby dla widza drażniąca, dlatego ustalono pewien schemat, który powtarzał się w każdej ze scen zawierającej wywiad. Każda scena rozpoczynała się od wprowadzenia lektora.

Następnym etapem montażu był wybór przebitek oraz odpowiednie umieszczenie ich w danej scenie. Przebitki w formie reportażowej służą przede wszystkim, jako uzupełnienie przeskoków między pociętymi fragmentami materiału oraz skrócenia wywiadów, w sposób niezauważalny dla widza.

Przedostatnim, ale również bardzo ważnym etapem montażu była korekcja koloru. W filmach reportażowych najważniejsza jest naturalność, dlatego ingerencja w obraz musi być praktycznie niewidoczna. Użyto więc zainstalowanych w programie filtrów, dzięki którym kolor obrazu został bardzo delikatnie ocieplony. Zmiany zabarwienia tonów można dokonywać w trzech podstawowych kanałach: dla tonów ciemnych (low), średnich (medium) oraz jasnych (high), co daje możliwość większej manipulacji kolorem i możliwość uzyskania ciekawych konfiguracji.

Ostatnim etapem montażu było wprowadzenie ujęć z napisami, animacji początkowej oraz belek, prezentujących nazwiska wypowiadających się osób. Te pierwsze miały za zadanie wprowadzenie widza do kolejnych scen, które nie można było ze sobą połączyć w żaden inny logiczny sposób, ponieważ film kolejno przedstawia praktyczne zastosowania technologii mowy. Do stworzenia napisów, animacji początkowej i belek wykorzystano dostępne wtyczki, które można zainstalować w programie. Należało jedynie z pośród ich ogromnej liczby wybrać te, które najlepiej wpasowywały się w klimat filmu.

Czasochłonność i skomplikowanie montażu obrazu jest nieodłącznie związane z gatunkiem produkcji. W przypadku realizowanego filmu popularno-naukowego proces ten okazał się dużo szybszy i prostszy, niż opisany w następnym rozdziale proces realizacji i montażu dźwięku.

## 4. Dźwięk w filmie

### 4.1. Techniki realizacji dźwięku na planie

Jeśli mówi się o dźwięku w filmie to należy rozpocząć od podstawowego podziału pracy z tym związanej. Aby film posiadał wszelakie walory dźwiękowe, należy najpierw zadbać o to, aby dźwięk na planie został odpowiednio nagrany, w zależności od wizji twórców. Dopiero po tym można mówić o takich procesach jak udźwiękawianie, nazywane również montażem dźwięku.

Technik realizacji dźwięku na planie jest tyle, ilu realizatorów dźwięku. Każdy ma inne metody, aby uzyskać jak najlepsze brzmienie dialogów, czy dźwięków z planu. Istnieją jednak od lat pewne niezmienniki w postaci sprzętu, z którym pracuje dźwiękowiec. Do podstawowych urządzeń wykorzystywanych podczas tego działania są oczywiście mikrofony. Najczęściej spotyka się mikrofony typu *shotgun*, czyli mikrofony pojemnościowe, posiadające hiperkardioidalną charakterystykę kierunkową. Używane są zazwyczaj do scen, w których aktor odgrywa monolog. Wtedy bardzo łatwo ustawić mikrofon tak, aby idealnie zbierał głos, który trzeba nagrać. Mikrofon ten, dzięki swojej budowie bardzo dobrze tłumi niechciane dźwięki tła, padające na membranę z innych, niż głos wybranej osoby, stron [Yewdall 2011, s. 75-76]. Nieodzownym elementem wspomagającym pracę tego typu mikrofonu na planie zdjęciowym jest rękojeść pistoletowa i obudowa typu Zeppelin, przedstawione na rysunku poniżej.



**Rys.3. Osłona typu Zeppelin i rękojeść pistoletowa**

*Źródło: www.dvcity.com*

Rękojeść ma za zadanie eliminację wibracji pochodzących z przypadkowych uderzeń w tyczkę, natomiast osłona ma chronić przed niepożądanymi podmuchami powietrza w mikrofon, które skutkują znacznymi zakłóceniami w nagraniu [Yewdall 2011, s. 84]. Mikrofony pojemnościowe są bardzo czułe, dlatego słyszalny jest nawet najmniejszy dotyk przeniesiony przez tyczkę. Ważne jest również, aby jeszcze przed nagraniem ocenić warunki akustyczne jakie panują na planie zdjęciowym, aby uniknąć zakłóceń. Na planach wykorzystuje się, chociaż znacznie rzadziej, mikrofony bezprzewodowe zwane powszechnie mikroportami. Mają one podstawową wadę, a mianowicie ciężko ukryć je w taki sposób, aby nie były widoczne w kamerze, a jednocześnie nie zbierały na przykład szelestu ubrania aktora. Z tego względu wykorzystuje się je tylko w sytuacjach, w których nie da się użyć mikrofonu na tyczce.

Kolejnym istotnym elementem ekwipunku dźwiękowca na planie jest mikser. W zależności od realizowanego projektu może to być urządzenie przenośne, posiadające od dwóch do sześciu kanałów, zasilane bateriami, lub duży mikser stacjonarny. Ten drugi nie różni się niczym od mikserów audio wykorzystywanych w studiach nagrań, posiadających korektory graficzne, tłumiki oraz mnóstwo gniazd umożliwiających podpięcie wszelakich procesorów efektowych. Z kolei dominującą cechą przenośnych mikserów audio jest ich poręczność i mobilność. Dlatego wyposażone są w kilka wejść mikrofonowych, wyjście stereo oraz pokrętła takich parametrów jak wzmocnienie i tłumienie danego kanału. Tego typu urządzenia są niezbędne podczas zgrywania dźwięku z planu, na którym używa się kilku mikrofonów.

Ostatni z elementów niezbędnych podczas realizacji dźwięku na planie jest rejestrator. Współcześnie wykorzystuje się już tylko cyfrowe rejestratory dźwięku, głównie ze względu na ich małe rozmiary, które znacznie ułatwiają pracę na planie, i niską cenę. W porównaniu ze stosowanymi wcześniej profesjonalnymi rejestratorami analogowymi, nie cechują się lepszą jakością nagranego materiału, jednak są wygodniejsze w obsłudze [Yewdall 2011, s. 66-67]. Do najpopularniejszych cyfrowych rejestratorów audio zaliczyć można te, produkowane przez firmę Zoom, Roland oraz Tascam. W zależności od modelu posiadają od 2 do 4 kanałów zewnętrznych mikrofonowych oraz wbudowany system stereo wykorzystujący zazwyczaj dwa mikrofony dookólne ustawione zgodnie z metodą nagrywania XY. Systemy te są także często wykorzystywane do nagrywania tła akustycznego z danej lokacji, czy charakterystycznych odgłosów, które zostaną później wykorzystane do udźwiękowania filmu.

## 4.2. Udźwiękowanie i metoda Foley'a

Montaż dźwięku ma dwie podstawowe funkcje: naprawczą i „upiększającą”. Pierwsza z nich jest niezbędna z punktu widzenia jakości prezentowanego w filmie dźwięku. Przy pracy nad nim ciągle zdarzają się niespójne fragmenty, przegapione głosy, czy niepożądane dźwięki nagrane na planie zdjęciowym. Na tym etapie usuwa się wszystkie niedociągłości eliminując je ze ścieżki, bądź też dogrywając brakujące fragmenty wypowiedzi aktorów. Druga funkcja montażu dźwięku ma za zadanie ubarwić tworzoną produkcję poprzez dodanie efektów, które nie występowały podczas realizacji materiału na planie. Mowa tu o specyficznych odgłosach, czy dźwiękach, których nie dało się nagrać, a mają znaczenie dla fabuły danego filmu.

Podczas montażu dźwięku można wyróżnić dwa podstawowe rodzaje scen: sceny dialogowe oraz sceny dźwiękowe. Sceny dźwiękowe to te, które nie zawierają dialogów, mogą natomiast posiadać głos narratora, czy efekty synchroniczne. W tego typu scenach zazwyczaj najpierw następuje montaż samego obrazu, a dopiero później dobiera się odpowiednie efekty dźwiękowe. Inaczej jest w przypadku scen dialogowych. Tutaj najistotniejszą rolę odgrywają wypowiedzi aktorów. Każde cięcie montażysty obrazu jest motywowane sensem dialogów i to on dokonuje ich edycji pod kątem czasu i wypowiedzanych treści. Montażysta dźwięku ma za zadanie jedynie sprawić, aby brzmienie mowy było wyraźne i posiadało odpowiednią barwę. Zdarza się, że po odpowiednim zmontowaniu obrazu jakość dialogu jest niezadowolająca i należy dokonać dubli dźwiękowych, czyli ponownego nagrania samego dialogu lub jego fragmentu w studio. Proces ten jest bardzo kosztowny i czasochłonny, aktor musi bowiem odtworzyć swoją kwestię tak, aby wyglądało to naturalnie. Z tego właśnie względu bardzo precyzyjnie planuje się nagrywanie dialogów na planie [Mascelli 2008, s. 263-264].

Najczęściej stosowanym podejściem podczas montażu dźwięku jest podział pracy od obróbki małych sekwencji do całkowitego zgrania. A więc pracę rozpoczyna się od miksowania każdej wypowiedzi aktora i każdego efektu dźwiękowego z osobna. Dokonuje się koniecznych operacji na pojedynczych sygnałach, co pozwala uniknąć większych błędów na późniejszych etapach pracy. Dopiero po doprowadzeniu do perfekcji pojedynczych efektów dokonuje się ostatecznego miksowania i zgrania [Mascelli 2008, s. 306-307].

Dużym ułatwieniem w procesie montażu dźwięku, a przede wszystkim w jego synchronizacji z obrazem stał się stworzony w 1997 roku format Open Media Framework, w skrócie OMF. Format ten służy do transferu plików między różnymi oprogramowaniami do obróbki plików audio i wideo. Zachowuje on jednocześnie wszelkie informacje m.in. na temat kodu czasowego, czy połączeń ze stosowanymi w innych programach efektami [10]. W praktyce pozwala to na przykład na wyeksportowanie z programu do montażu elektronicznego sekwencji wszystkich ścieżek dźwiękowych, odpowiednio ułożonych i pociętych pod kątem filmu. Plik taki może zostać zaimportowany do platformy DAW (Digital Audio Workstation) i tam odpowiednio obrabiany przez montażystę dźwięku, bez naruszania kodów czasowych, a więc relacji z obrazem. Współcześnie wszystkie najnowsze wersje platform DAW takich jak Pro Tools, Logic, Cubase, Samplitude obsługują ten format, dzięki czemu proces montażu dźwięku jest znacznie usprawniony. Nie dotyczy się to tylko transferu dźwięku, ale również plików wideo.

Mówiąc o udźwiękowieniu filmów nie można zapomnieć o człowieku, który wprowadził w życie najczęściej stosowaną metodę uzyskiwania efektów synchronicznych – Jacku Foley’u. Był on odpowiedzialny za pierwsze udźwiękowienia filmów pod koniec lat 20-tych ubiegłego wieku. Wtedy to w studiu wytwórni Fox, osoby odpowiedzialne za efekty dźwiękowe przyglądała się projekcji filmu i w odpowiednich momentach tworzyły lub wydawały odgłosy pasujące do danego ujęcia. Wszystko to było w sposób ciągły nagrywane na taśmę i tak tworzono ścieżkę dźwiękową do danej produkcji. Dzięki późniejszemu rozwojowi optycznej realizacji dźwięku możliwe było odpowiednie cięcie efektów i jeszcze lepsze ich dopasowywanie [Yewdall 2011, s. 418]. Współcześnie metoda ta wykorzystywana jest w różnych gatunkach i stylach filmowych, animacjach i grach komputerowych. Sam sposób tworzenia efektów synchronicznych nie zmienił się przez te lata. Nadal wykorzystuje się różne przyrządy, które imitują realne odgłosy oraz tworzy się podłogi z różnego materiału, aby można było kreować odpowiednie kroki. Różnica polega jedynie na tym, że nie nagrywa się już na taśmach. Efekty są bezpośrednio nagrywane na stację roboczą (DAW) i tam odpowiednio edytowane i miksowane.

### 4.3. Realizacja i montaż dźwięku w filmie *Technologie mowy*

Pierwszym etapem prac nad dźwiękiem, w tworzonym na potrzeby pracy filmie, była realizacja dźwięku na planie zdjęciowym. Do tego celu wykorzystano mikrofon Sony ECM-673, który posiada przetwornik pojemnościowy, kierunkową charakterystykę superkardioidalną oraz duży zakres przenoszonych częstotliwości. Tego typu mikrofony są najczęściej stosowane na planach zdjęciowych, ponieważ w porównaniu do mikrofonów dynamicznych lepiej radzą sobie z przenoszeniem wysokich częstotliwości oraz szybkich impulsów. Konieczna jest również superkierunkowa charakterystyka, która umożliwia precyzyjne zbieranie dźwięku nawet z dalszych odległości. Możliwe jest to dzięki temu, że osie najmniejszej skuteczności, czyli zależności między napięciem na wyjściu mikrofonu a kątem padania fali dźwiękowej, znajdują się z tyłu po bokach w odległości ok.  $120^\circ$ . Dodatkowo wykorzystywany podczas nagrań na planie mikrofon charakteryzuje się dużą czułością i niskim poziomem szumów własnych na poziomie mniejszym niż 17dB SPL, co gwarantuje zrozumiałość mowy, nawet przy wysokim poziomie tła akustycznego [11]. Jest to niezwykle ważna cecha szczególnie podczas nagrań różnego typu reportażu w terenie, kiedy często nie ma możliwości zapanowania nad tym, co dzieje się dookoła planu zdjęciowego, a nagranie jest w danym miejscu nieuniknione. W przetworniku wbudowany jest także filtr górnoprzepustowy, który również wzmacnia efekt zmniejszenia skutków niepożądanych szumów otoczenia.

Ze względu na bardziej reportażową formę projektu, podczas realizacji na planie dźwięk był nagrywany bezpośrednio na kamerę. Tego typu zabieg pozwala na dużą oszczędność czasu podczas montażu materiału, dzięki braku konieczności synchronizacji obrazu z dźwiękiem. Współcześnie produkowane kamery posiadają już parametry rejestracji dźwięku porównywalne do przenośnych rejestratorów cyfrowych audio, dlatego też na tym etapie nie było konieczności wykorzystania dodatkowego urządzenia.

Kolejnym etapem prac nad dźwiękiem w projekcie były nagrania lektorskie. Tutaj uzyskanie profesjonalnego efektu sprawiło nieco więcej problemów, ze względu na brak dostępu do profesjonalnej kabiny lektorskiej. Dlatego więc na potrzeby nagrań stworzono bardzo prowizoryczną kabinę, która miała zapewnić jak najmniejszy czas pogłosu. Do tego celu wykorzystano kilka materacy, oraz kotary zawieszane w pewnej odległości od ścian, co pozwoliło uzyskać większe tłumienie dźwięku przynajmniej w wyższym zakresie częstotliwości mowy ludzkiej. Efekt końcowy okazał się zaskakująco dobry jak na warunki



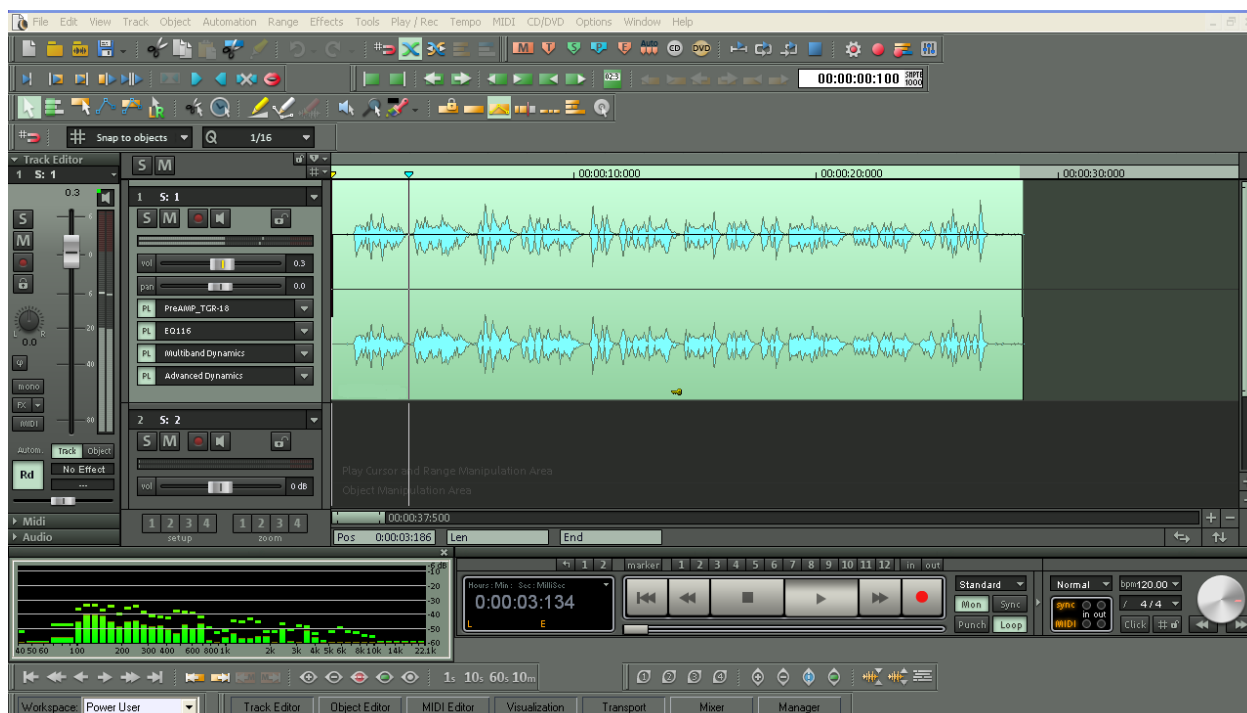
w jakich pracował lektor. Okazało się bowiem, że kotary i materace skutecznie zminimalizowały czas pogłosu w najbardziej pożądanym zakresie częstotliwości, odpowiadającemu kobiecemu głosowi. Do nagrań wykorzystano pojemnościowy mikrofon wokalny AKG C5 oraz cyfrowy, przenośny rejestrator audio Roland R-26. Pierwsze urządzenie posiadało kardoidalną charakterystykę przenoszenia, która gwarantuje dużą czułość mikrofonu. Maksymalna dynamika to aż 140 dB, dlatego też mikrofon ma częste zastosowanie estradowe. W związku z wykorzystanym w urządzeniu przetwornikiem potrzebowało ono zewnętrznego zasilania fantomowego, które zapewniał cyfrowy rejestrator audio, na którym w formacie *wave* z częstotliwością próbkowania 48 kHz odbywał się 24-bitowy zapis. Zdecydowano się właśnie na takie parametry, ponieważ są one zgodne z tymi zapisywanymi przez kamerę podczas nagrań. W ten sposób zapewniono szybszy montaż dźwięku, niewymagający między innymi operacji *resamplingu* otrzymanego sygnału. W celu uzyskania dużego nasycenia brzmienia w niskich częstotliwościach postanowiono umieścić mikrofon tak blisko ust lektora, jak tylko było to możliwe. W mikrofonach kierunkowych (ciśnieniowo-gradientowych) występuje wtedy efekt zwany zbliżeniowym, który sprawia że wraz ze zmniejszeniem się odległości między ustami osoby mówiącej a membraną wzrasta „uwypuklenie” niskich częstotliwości. Efekt ten jest wręcz permanentnie wykorzystywany w radio. Istnieje jednak pewna znacząca wada takiej techniki nagrań. Występuje wtedy większa eksploatacja spółgłosek zwanych wybuchowymi, czyli np. *p, t, b*. Aby temu zapobiec również w tych konkretnych nagraniach lektora wykorzystano *pop filter*, który jest filtrem mechanicznym pozwalającym na eliminację szybko poruszającego się powietrza odpowiedzialnego za ową wybuchowość.

Jak już wspomniano wcześniej nagrania były bezpośrednio zgrywane i zapisywane na przenośnym rejestratorze audio firmy Roland. Urządzenie to posiada dwa typy wbudowanych mikrofonów stereo o charakterystyce dookólnej oraz kierunkowej, a także parę wejść combo XLR/TRS. W nagraniu potrzebne było jedynie jedno wejście XLR, za pomocą którego podłączony został mikrofon. Rejestrator został też wyposażony w 24-bitowy przetwornik analogowo-cyfrowy i cyfrowo-analogowy oraz możliwość próbkowania nawet z maksymalną częstotliwością 96kHz [12]. W czasie nagrań wystąpił jednak problem najprawdopodobniej związany ze złym dopasowaniem impedancji mikrofonu, ponieważ sygnał rejestrował się na sięgającym chwilami poziomie -36 dB. W związku z brakiem możliwości skorzystania z innego mikrofonu, rejestratora, bądź użycia zewnętrznego wzmacniacza mikrofonowego, jedynym rozwiązaniem było możliwie jak największe wzmocnienie sygnału wejściowego.

Spowodowało to jednak znaczny wzrost szumów własnych mikrofonu słyszanych w nagraniu. Na tym etapie prac nie uznano tego typu zakłóceń za zbyt szkodliwe. Jednak jak okazało się na etapie edycji, szумы stanowiły dość duży problem.

Po zakończeniu realizacji dźwięku na planie oraz nagrań lektorskich nastąpił drugi etap prac związanych z montażem dźwięku, czyli edycja i miks, stanowiąca w przypadku filmu reportażowego, podstawowe elementy udźwiękowania produkcji. Nie ma bowiem potrzeby dodawać większej ilości efektów dźwiękowych, poza muzyką stanowiącą podkład. Jedynym udźwiękowanym fragmentem w przypadku realizowanego projektu były napisy, które wprowadzały widza w następne fragmenty filmu. W tym przypadku, ze względu na wykorzystanie dość charakterystycznej animacji postanowiono na dodanie efektów dźwiękowych, które jednocześnie komponowałyby się z następującą po nich muzyką stanowiącą niejako tło filmu.

Głównym narzędziem wykorzystywanym do montażu dźwięku w projekcie była stacja robocza Samplitude w wersji 11. Program jest profesjonalną platformą DAW, która zapewnia możliwość aranżacji, edycji, miksowania oraz masteringu audio. Poniżej zamieszczono zdjęcie głównego okna do pracy w programie.

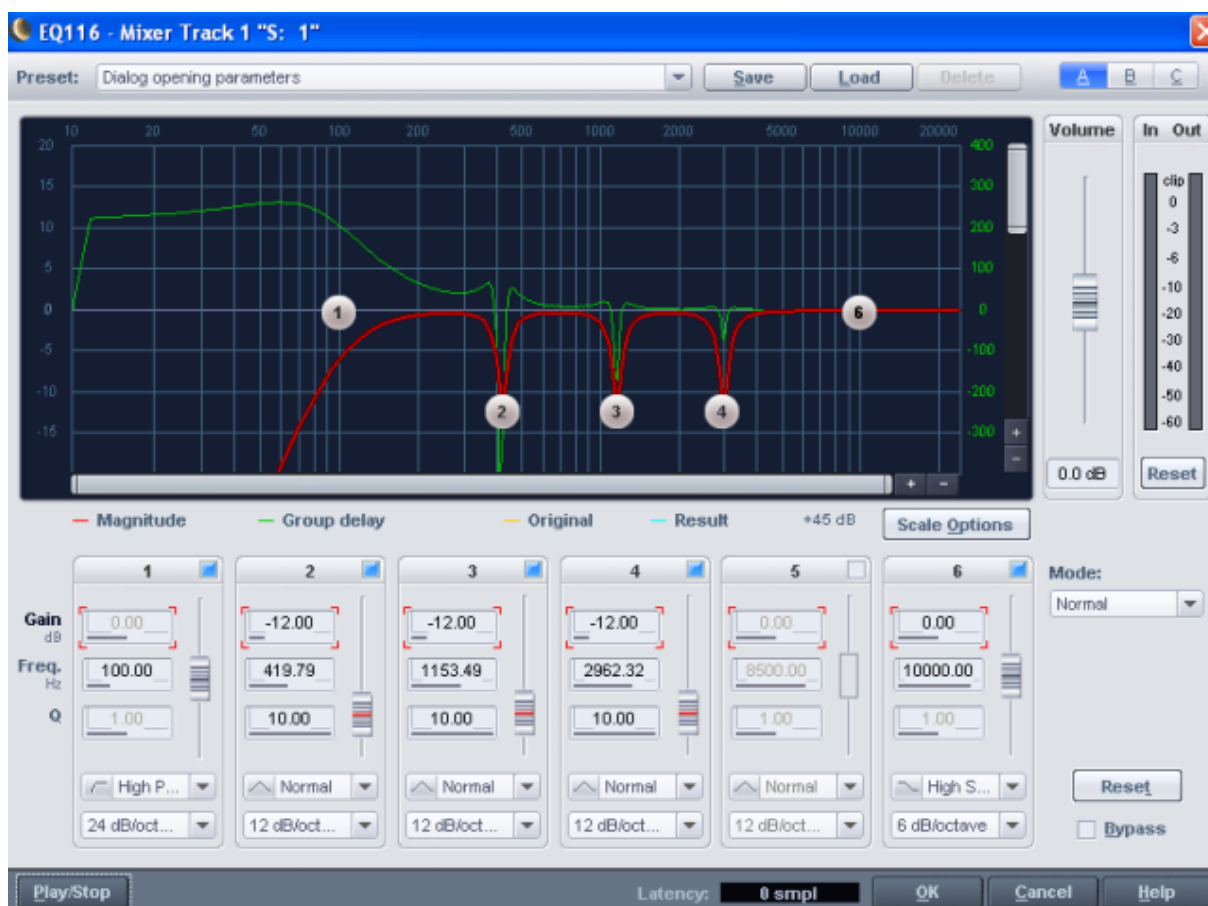


**Rys. 4. Okno pracy programu Samplitude**

*Źródło: opracowanie własne*

Ustawienia pracy w programie takie jak rodzaj okna, wykresów i parametrów dźwięku, skrótów klawiszowych, można dobrać dowolnie. Stacja robocza umożliwia również wpięcie do jednej ze ścieżek pliku wideo, dzięki czemu możliwe jest udźwiękowanie wcześniej zmontowanego filmu. Możliwość taka sprawia, że platformy można używać do profesjonalnego udźwiękowania filmów, bowiem możliwe jest perfekcyjne dopasowanie efektu dźwiękowego do danej klatki filmu. Większość programów do obróbki wideo posiada możliwość edycji ścieżek audio jednak w bardzo podstawowym zakresie. Tutaj mamy dostęp do różnego rodzaju procesorów efektowych, które pozwalają na nieograniczoną obróbkę dźwięku.

Edycja i miks nagrań lektorskich stanowiły niejako osobną część pracy nad montowanym dźwiękiem. Założenie bowiem było takie, aby uzyskać barwę głosu, która jest charakterystyczna dla lektorów radiowych i telewizyjnych. Do tego celu, po złożeniu ostatecznych wersji wypowiedzi, wykorzystano trzy podstawowe narzędzia, które są zainstalowane na praktycznie każdej platformie DAW, również na tej używanej w projekcie. Mowa tu o *equalizerze*, kompresorze pasmowym oraz limiterze. Podstawowym zadaniem tego pierwszego, była korekcja częstotliwościowa sygnału. Użyto do tego zainstalowanego na stacji roboczej korektora parametrycznego, w którym możliwa jest regulacja częstotliwości środkowej, szerokości pasma oraz wzmocnienia. Podstawowa zasada podczas stosowania korektorów zarówno graficznych, jak i parametrycznych, to osłabianie a nie wzmacnianie dźwięku. A więc częściej eliminuję się pasma, w których występują rezonanse lub przesłuchy, niż podbija pasma, które nam odpowiadają. Takie działanie zapobiega przesterowaniu sygnału. Zasadę tę stosowano również podczas korekty częstotliwościowej nagrań lektorskich wykorzystywanych w filmie. Najpierw na sygnał nałożono filtr górnoprzepustowy o częstotliwości odcięcia ok.100 Hz, a następnie za pomocą filtru o bardzo dużej dobroci ( $Q=10$ ) i wzmocnieniu (12dB) przeszukano sygnał pod kątem słyszalnych rezonansów i przesłuchów. Znalaziono kilka częstotliwości rezonansowych, głównie w paśmie 500-2500Hz, które następnie ściszo, używając tłumika ustawionego na wartość -12dB. W ten sposób sygnał stał się bardziej wyrównany w całym paśmie częstotliwości i przygotowany do użycia kompresora. Poniżej zaprezentowano główne okno zastosowanego *equalizera*.

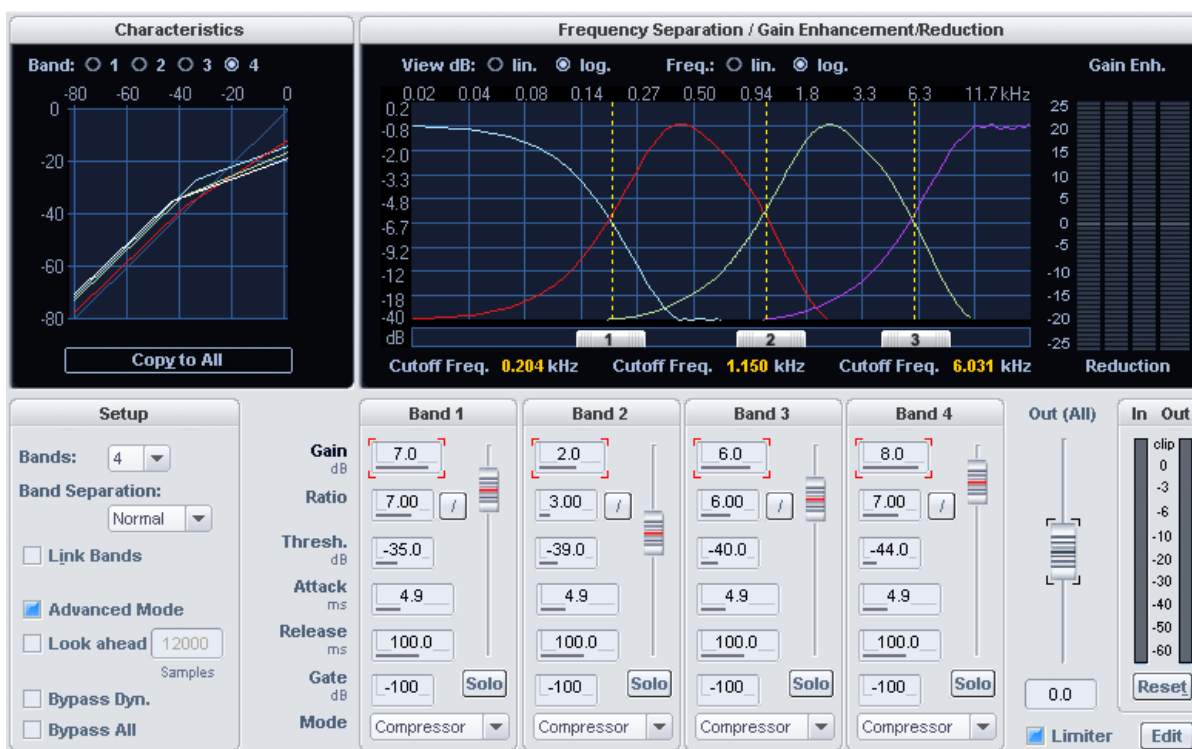


**Rys. 5. Equalizer wykorzystywany do korekty częstotliwościowej**  
*Źródło: opracowanie własne*

W celu nadania sygnałowi mowy specyficznej radiowej barwy niezbędne jest użycie kompresora pasmowego. Kompresor to procesor dynamiczny, który ma za zadanie zmniejszenie różnicy pomiędzy najcichszymi i najgłośniejszymi fragmentami sygnału. Posiada on pięć podstawowych parametrów, które opisują jego pracę i wpływają na barwę dźwięku. Należą do nich [Sztekmler 2003, s. 65-67]:

- threshold – wielkość napięcia powyżej której następuje redukcja sygnału,
- ratio – współczynnik kompresji, czyli zależność między wielkością napięcia wejściowego i wyjściowego powyżej progu threshold,
- attack – czas po jakim zadziała procesor i nastąpi kompresja,
- release – czas powrotu do wzmocnienia przed zadziałaniem kompresora,
- output gain – regulator napięcia wyjściowego, służący to podniesienia poziomu sygnału, który uległ kompresji.

Zaprezentowane parametry są bardzo istotne z punktu widzenia barwy dźwięku i wszystkie zostały eksperymentalnie dobrane do kompresji mowy lektora. W przypadku chęci uzyskania charakterystycznej barwy radiowej ważne jest wykorzystanie kompresora pasmowego, który pozwala na osobną redukcję sygnału w zadanych pasmach częstotliwościowych. W realizowanym projekcie wykorzystano zainstalowany w programie Samplitude czteropasmowy kompresor. Poniżej zaprezentowano rysunek przedstawiający zastosowane wartości parametrów w kolejnych pasmach częstotliwościowych.



**Rys. 6. Parametry wykorzystanego kompresora pasmowego**

*Źródło: opracowanie własne*

Parametry *attack* i *release* zostały ustawione tak samo dla wszystkich pasm, aby poziom dźwięku był wyrównany dla wszystkich częstotliwości. Czas ataku jest bardzo krótki (ok. 5ms), aby nie przepuścić żadnych szybkich impulsów sygnału. Czas powrotu musi być natomiast wybierany eksperymentalnie, w zależności od kompresora, powinien on być bowiem w tym przypadku jak najkrótszy, ale również na tyle długi, żeby nie generował żadnych trzasków. Często zdarza się, że zbyt mała wartość parametru *release* sprawia, że w nagraniu słyszalne są dodatkowe zniekształcenia. Jeśli mowa o doborze pozostałych parametrów, to początkowo ustawiony został współczynnik kompresji, w taki sposób aby najmocniej zredukowane zostało pasmo dolne i górne, w mniejszym stopniu wyższe

środkowe, a tylko delikatnie niższe środkowe. Duża wartość parametru *ratio* powoduje, że w danym paśmie słyszalne są bardzo ciche a charakterystyczne dźwięki, które dodają głębi głosu w niższych częstotliwościach i zrozumiałości w środkowych wyższych. Pasma zawierające częstotliwości środkowe niższe jest bardzo mało kompresowane, ponieważ z punktu widzenia zabawienia głosu lektorskiego jest bardzo mało ciekawe, a nie odpowiada za zrozumiałość mowy. Parametr *threshold* był dobierany w taki sposób, aby redukcja w danym paśmie wahała się od 3 do 6 dB. Po dokonaniu tych operacji należało wzmocnić zredukowany sygnał przynajmniej do poziomu zerowego. Aby jednak barwa nabrała jeszcze większego radiowego charakteru jedynie pasmo niższe środkowe wzmocniono do poziomu zero, a resztę podniesiono o kilka decybeli ponad ten poziom. Należy tutaj jeszcze wspomnieć o poziomie szumów, o którym mówiono podczas omawiania samej realizacji nagrania. Zapisane wówczas zniekształcenia zostały dość słyszalnie wyeksponowane przez kompresję, co niestety było błędem realizacji, którego nie udało się naprawić w procesie montażu. Do zagłuszenia tych zniekształceń posłużyła muzyka, która tworzyła tło do całego filmu. Jedynie w ten sposób można było nadać głosowi lektora odpowiedniej barwy, bez wyraźnie słyszalnych szumów.

Ostatnim procesorem, wykorzystanym podczas edycji głosu lektora był limiter. Urządzenie to co do zasady działa tak samo jak kompresor, czyli redukuje sygnał, który przekroczy pewną ustaloną granicę. Różnica jest taka, że współczynnik kompresji nie jest dowolnie ustawiany, a bliski nieskończoności. Podobnie jest z czasem włączenia kompresora, który jest bliski zeru. W tym przypadku limiter włączony na końcu toru pozwala na jeszcze większe wydobycie brzmienia uzyskanego za pomocą kompresora. Należy jednak pamiętać, że każdy procesor powoduje zniekształcenia i należy dobrać jego wartość „na słuch”. W projekcie poziom limitera został ustawiony na -8 dB.

Proces montażu dźwięku był z pewnością najdłuższym etapem podczas tworzenia filmu, głównie ze względu na nagrania lektorskie, które były konieczne dla jego spójności. Ostateczny efekt zdaniem autora jest dobry a założone cele związane z jakością dźwięku zostały zrealizowane.

## 5. Zakończenie

Prezentowana praca inżynierska składała się z części projektowej i opisowej. Głównym celem projektowej części było stworzenie krótkometrażowego filmu popularno-naukowego prezentującego systemy, realizowane przez Zespół Przetwarzania Sygnałów na Akademii Górniczo-Hutniczej, związane z technologiami mowy. Założeniem było, aby treści filmu były zrozumiałe dla osób, które nie posiadają szczegółowej wiedzy technicznej z tego zakresu. Po ukończeniu produkcji można stwierdzić, że to założenie zostało osiągnięte.

W opisowej części pracy skupiono się natomiast na opisaniu realizacji projektu zarówno od strony technicznej, jak i artystycznej. Zaprezentowane zostały wymagania i podstawowe metody tworzenia kolejnych etapów filmu, na podstawie literatury przedmiotu. Opisując kolejne procesy tworzenia produkcji prezentowano najpierw historię związaną z daną dziedziną, a następnie współczesne metody. Autor przedstawił również swoje rozwiązania dotyczące realizacji i montażu dźwięku, montażu obrazu oraz tworzenia scenariusza i realizacji zdjęć.

Wszystkie założone cele pracy zostały zrealizowane, a końcowy efekt, jakim jest krótkometrażowy film popularno-naukowy, jest dostępny na kanale TV AGH platformy *YouTube*, pod tytułem: Technologie Mowy.

## Bibliografia:

1. Ament V., *The Foley grail. The art. Of performing sound for film, games and animation*, wyd. Focal Press, USA 2009.
2. Aijon D., *Gramatyka języka filmowego*, Wydawnictwo Wojciech Marzec, Warszawa 2010.
3. Helman A., *O dziele filmowym*, Wydawnictwo Literackie, Kraków 1981.
4. Kingdon T., *Sztuka reżyserii filmowej*, Wydawnictwo Wojciech Marzec, Warszawa 2007.
5. Mascelli J. V., *5 tajemnic warsztatu filmowego*, Wydawnictwo Wojciech Marzec, Warszawa 2008.
6. Murch W., *W mgnieniu oka. Sztuka montażu filmowego*, Wydawnictwo Wojciech Marzec, Warszawa 2006.
7. Yewdall D. L., *Dźwięk w filmie. Teoria i praktyka*, Wydawnictwo Wojciech Marzec, Warszawa 2011.
8. [www.moviola.com/history](http://www.moviola.com/history) [data dostępu: 27 grudnia 2012]
9. Hendykowski M., *Język ruchomych obrazów*, wyd. ARS NOVA, Poznań 1999.
10. [www.edlmax.com/FormatOmf](http://www.edlmax.com/FormatOmf) [data dostępu: 28 grudnia 2012]
11. [www.sony.pl/pro/lang/en/pl/product/shotgun/ecm-673](http://www.sony.pl/pro/lang/en/pl/product/shotgun/ecm-673) [data dostępu: 28 grudnia 2012]
12. [www.rolandpolska.pl](http://www.rolandpolska.pl) [data dostępu: 28 grudnia 2012]
13. Sztekmler K., *Podstawy nagłośnienia i realizacji nagrań*, wyd. Narodowe Centrum Kultury, Warszawa 2003.
14. [pl.wikipedia.org/wiki/Bracia\\_Lumiere](http://pl.wikipedia.org/wiki/Bracia_Lumiere) [data dostępu: 25 grudnia 2012]
15. Ober K., *Proces digitalizacji*, [lib.psnc.pl/Content/105/Proces+digitalizacji.pdf](http://lib.psnc.pl/Content/105/Proces+digitalizacji.pdf) [data dostępu: 3 stycznia 2013]
16. [www.pro.sony.com](http://www.pro.sony.com) [data dostępu: 3 stycznia 2013]
17. Fernsham R., *Jak napisać scenariusz*, Wydawnictwo Literackie, Kraków 1998.



## Spis rysunków:

Rys. 1. Półzblizenie	10
Rys. 2. Okno pracy programu Sony Vegas	18
Rys.3. Osłona typu Zeppelin i rękojeść pistoletowa	20
Rys. 4. Okno pracy programu Samplitude	26
Rys. 5. <i>Equalizer</i> wykorzystywany do korekty częstotliwościowej	28
Rys. 6. Parametry wykorzystanego kompresora pasmowego	29