

WPLYW PARAMETRÓW KOMPRESJI DYNAMICZNEJ NA SUBIEKTYWNĄ GŁOŚNOŚĆ NAGRANIA INSTRUMENTU MUZYCZNEGO

Karol CZESAK¹

Piotr KLECZKOWSKI²

¹Akademia Górniczo Hutnicza im. Stanisława Staszica w Krakowie, 30-059 Kraków,
al. Mickiewicza 30, czesak@agh.edu.pl

²Akademia Górniczo Hutnicza im. Stanisława Staszica w Krakowie, 30-059 Kraków,
al. Mickiewicza 30, kleckow@agh.edu.pl

W produkcji muzycznej zauważamy zwiększającą się tendencję do zawężania zakresu dynamicznego nagrań muzycznych. Głównym założeniem jest tu zwiększenie całościowej głośności nagrania. W zdecydowanej większości przypadków, pojęcie głośności odnosi się do wartości skutecznej sygnału w kanale transmisji. W rezultacie, materiał często jest obrabiany przy użyciu agresywnej kompresji dynamicznej. Jak pokazuje doświadczenie inżynierów dźwięku – subiektywne odczucie głośności nie jest warunkowane wyłącznie przez stopień kompresji, ale też przez inne jej parametry. Celem tej pracy jest zbadanie, które parametry kompresji dynamicznej wywierają istotny wpływ na subiektywnie postrzeganą głośność zróżnicowanego materiału muzycznego. Stwierdzono, że jest możliwe takie dobranie parametrów kompresji dynamicznej, by nagranie sprawiało wrażenie głośnego, bez stosowania nadmiernego stopnia kompresji.

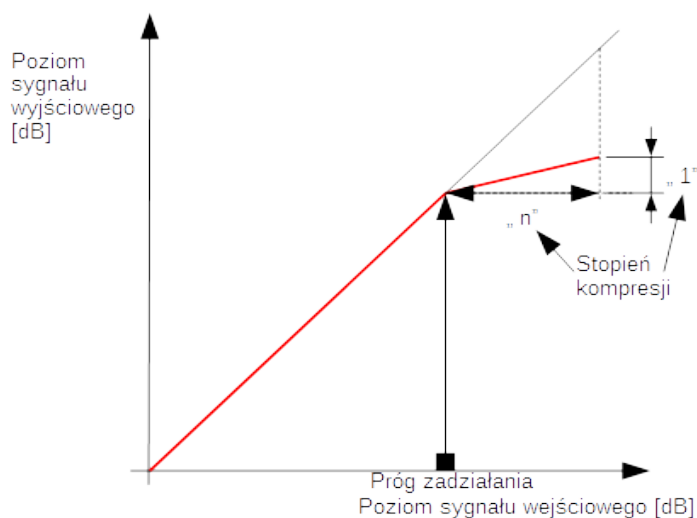
1. WSTĘP

Kompresja dynamiki jest procesem mającym na celu zredukowanie zakresu dynamiki materiału dźwiękowego, poprzez redukcję wzmocnienia sygnału, po przekroczeniu przezeń założonego progu. Pozwala ona zabezpieczyć kanał transmisji przed przesterowaniem, a także zwiększyć komfort odsłuchu materiału. Co więcej – już w latach 50tych, na podstawie obserwacji użytkowników szaf grających, stwierdzono, że materiał o większym poziomie skutecznym jest postrzegany przez odbiorcę jako atrakcyjniejszy. Prowadziło to do sukcesywnego zwiększania poziomu skutecznego nagrań na przestrzeni kolejnych dziesięcioleci. W latach 90tych XX wieku i u progu XXI stulecia, możemy mówić wręcz o „wojnie głośności” [1]. Na ową „Wojnę głośności” postanowiły odpowiedzieć Międzynarodowa Organizacja Telekomunikacyjna (ITU), oraz Europejska Unia Nadawców (EBU), publikując liczne dokumenty, mające na celu uzgodnienie poziomu głośności sygnału w mediach elektronicznych, w rozumieniu jej zakresu dynamicznego oraz wartości skutecznej. W związku z powyższym, zdajemy sobie pytanie: „Czy jest możliwe wyprodukowanie materiału dźwiękowego, spełniającego przyjęte zalecenia techniczne dotyczące głośności, jednak percepcyjnie głośniejszego od innych, sobie podobnych?”. Próba znalezienia odpowiedzi na tak postawione pytanie, okazała się być motywacją do znalezienia zależności między parametrami kompresji dynamicznej, a subiektywnym odczuciem głośności i atrakcyjności materiału, bazującym na podstawie komfortu i poczucia estetyki słuchacza.

2. ROZPATRYWANE PARAMETRY KOMPRESJI DYNAMICZNEJ

Kompresor dynamiki jest narzędziem przekształcającym sygnał w sposób nieliniowy, działającym w dziedzinie czasu. Wspomniane nieliniowości zachodzą jednak w stanach przejściowych, co pozwala nam zaliczyć kompresor do narzędzi liniowych, zmiennych w czasie [2]. Kompresor dynamiczny jest procesorem typu adaptatywnego – wprowadzane przezeń zmiany są zależne od sygnału podawanego na jego wejście, za którym to sygnałem procesor ten niejako „podąża”, w wyniku czego, jego stosowanie pozostaje nie bez wpływu na barwę dźwięku, a niekiedy wprowadza też – mniej lub bardziej słyszalne artefakty. Dzięki tej właściwości – kompresor może być stosowany nie tylko jako zabezpieczenie materiału przed wystąpieniem zniekształceń, ale też jako narzędzie kreacji brzmienia.

Charakterystykę statyczną kompresora definiują parametry: próg zadziałania i stopień kompresji. Próg zadziałania wyrażany jest w skali decybelowej, gdzie poziomem odniesienia jest nominalny poziom sygnału w kanale transmisji. Próg zadziałania parametrem określającym poziom, po przekroczeniu którego przez sygnał podany na wejście kompresora, zaczyna zachodzić tłumienie tegoż sygnału. Stopień kompresji wyrażany jest poprzez iloraz „n:1”, definiujący stosunek „nadwyżki” sygnału wejściowego, przekraczającego wartość progową, do „naddatku” sygnału wyjściowego pozostającego powyżej tejże wartości. Zmienna „n” zawiera się tutaj w przedziale (1: +∞). Charakterystyka statyczna teoretycznego modelu kompresora z zaznaczonymi opisanymi powyżej parametrami przedstawiona została na rysunku 1.



Rys. 1. Charakterystyka statyczna kompresora

Poziom, o jaki sygnał na wyjściu ulega redukcji w stosunku do sygnału na wejściu, określony jest wzorem:

$$S = (X - P) * \left(1 - \frac{1}{n}\right) \quad (1)$$

gdzie:

S – założony poziom redukcji wzmocnienia [dB]

X – poziom sygnału wejściowego [dB]

P – próg zadziałania kompresora [dB]

n – zmienna determinująca stopień kompresji postaci: n:1

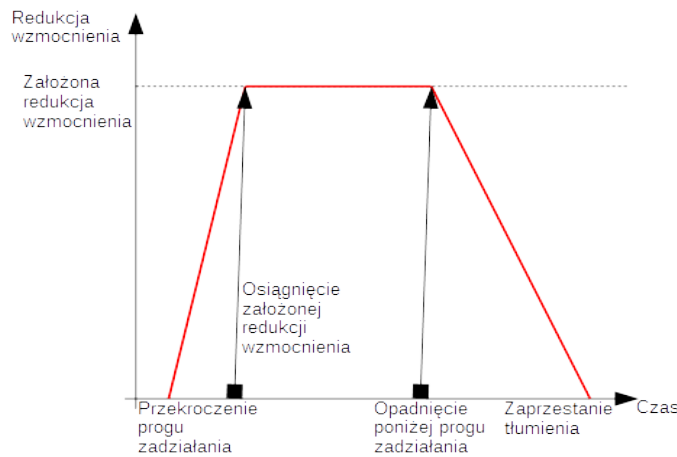
Co pozwala nam określić sygnał wyjściowy Y jako:

$$Y = X - S \quad (2)$$

W związku z powyższym, stopień kompresji przybiera postać:

$$R = \frac{X - P}{Y - P} \quad (3)$$

W celu uniknięcia artefaktów, związanych z rkompresją dynamiki, przebieg redukcji wzmocnienia charakteryzuje się pewną obwiednią czasową. Historycznie, wynikała ona ze swoistej bezwładności układów analogowych implementujących proces kompresji dynamiki. Obecnie jednak, obwiednia ta opisana jest czasem ataku oraz czasem opadania. Obwiednia ta – dla skali decybelowej, przedstawiona została na rysunku 2.



Rys. 2. Obwiednia czasowa tłumienia sygnału

Pamiętajmy jednak, iż kompresor w rzeczywistości pracuje na sygnale napięciowym, mającym charakter liniowy, toteż w rzeczywistości czas ataku przyjmuje postać:

$$t_{at} = 1 - e^{-2,2 * T_s / T_{at}} \quad (4)$$

gdzie:

t_{at} – rzeczywisty czas ataku kompresora [s]
 T_s – okres próbkowania [s]
 T_{at} – czas ataku w skali decybelowej [s]

Analogiczną postać przyjmuje przeliczenie czasu zwolnienia:

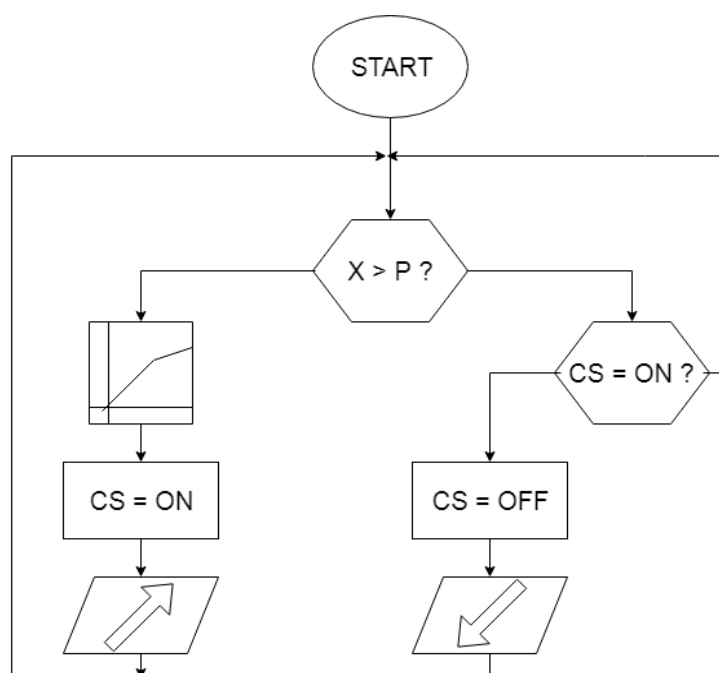
$$t_{op} = 1 - e^{-2,2 * T_s / T_{op}} \quad (5)$$

gdzie:

t_{op} – rzeczywisty czas opadania kompresora [s]
 T_s – okres próbkowania [s]
 T_{op} – czas opadania w skali decybelowej [s]

Wprowadzenie regulacji czasu ataku i zwolnienia kompresora nie tylko umożliwia ustrzeżenie się przed potencjalnymi artefaktami i pozwala lepiej dostosować własności kompresji do rodzaju obrabianego materiału, ale też stwarza możliwość zastosowania kompresora w charakterze narzędzia do kształtowania obwiedni amplitudowej sygnału [3]. Przykładem takiego zastosowania może być kontrola transjentów;

Zgodnie z powyższymi rozważaniami, oznaczając stan pracy kompresora zmienną CS , przyjmując wartości „prawda” lub „fałsz”, działanie kompresora, za pomocą schematu blokowego, przedstawione zostało na rysunku 3.



Rys. 3. Schemat blokowy kompresora dynamiki

W celu uzyskania kompresora, dokładnie spełniającego powyższe założenia, konieczna była jego samodzielna implementacja programowa przez autora. Dla łatwości użytkowania, został on zaimplementowany w postaci wtyczki VST, przy użyciu środowiska JUCE. Widok graficznego Interfejsu użytkownika przedstawia rysunek 4.



Rys. 4. Graficzny interfejs użytkownika kompresora, uruchomionego wewnątrz oprogramowania Cockos Reaper

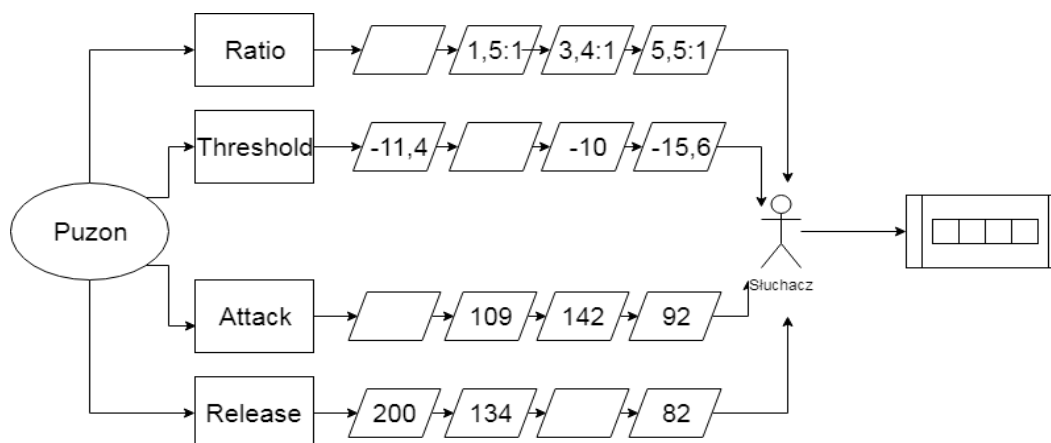
3. DOBÓR PRÓBEK I TESTY ODSŁUCHOWE

Analizę parametrów kompresji na subiektywnie odbieraną głośność zdecydowano się przeprowadzić dla sześciu instrumentów, chętnie poddawanych przez realizatorów dźwięku kompresji dynamicznej. Były to: Gitara basowa, puzon, saksofon, skrzypce, werbel i wokal. Próbki te pochodziły z wielośladowego nagrania utworu kompozycji Andrzeja Zielińskiego do słów Agnieszki Osieckiej „Nie całuj mnie pierwsza”, w bigbandowej aranżacji Wojciecha Zwierniaka.

Dla każdego z parametrów, dana próbka prezentowana była w testach odsłuchowych czterokrotnie: z trzema różnymi wartościami analizowanego parametru oraz w postaci nieskompresowanej. Wartości parametrów zostały dobrane doświadczalnie. Następnie – wartość skuteczna wszystkich próbek została znormalizowana do poziomu -30LKFS. Każda z próbek trwała od 12 do 15 sekund.

Słuchacz dostawał do odtworzenia zestaw plików, z których każdy zawierał cztery próbki brzmienia instrumentu. Gdy emitowane były próbki ze zmieniającym się konkretnym parametrem kompresji, wszystkie pozostałe parametry pozostawały niezmiennie. Kolejność prezentowania próbek w obrębie danej grupy była określona w sposób pseudolosowy. Zadaniem słuchacza było uszeregowanie – dla każdego z plików – próbek od uważanych przezeń za najcichsze, do uważanych przezeń za najgłośniejsze. Łącznie – słuchacz otrzymał do oceny 24 zestawy próbek. Przebieg testu obrazuje rysunek 5.

Analizie statystycznej zostały poddane wyniki uzyskane z przebadania 15 słuchaczy. Głównie były to osoby młode - Studenci kierunku Inżynieria Akustyczna na Krakowskiej AGH oraz młodzież korzystająca z oferty kulturalnej Świetlicy Wiejskiej w Gierczycach. Wszyscy respondenci pytani byli o wiek oraz doświadczenie w pracy z dźwiękiem lub grze na instrumencie. Dało się zauważyć, że osoby „grające” zgłaszają inne uwagi i spostrzeżenia dotyczące prowadzonego badania, co pozwala wnioskować, że funkcjonują u nich inne mechanizmy percepcji głośności dźwięku. Statystyczna istotność uzyskiwanych wyników oceniana była za pomocą testów Kruskala – Wallisa i wielokrotnego porównania średnich.



Rys. 5. Schemat przebiegu testów odsłuchowych

4. PREZENTACJA UZYSKANYCH WYNIKÓW

W przypadku każdego z badanych parametrów kompresji dynamicznej uzyskano statystycznie istotne wyniki, dla co najmniej dwu instrumentów. Najwięcej – bo aż dla pięciu

instrumentów – statystycznie istotnych różnic w percypowanej głośności uzyskano dla progu zadziałania. W przypadku stopnia kompresji – statystycznie istotne różnice pojawiły się dla czterech instrumentów. Dla parametrów określających obwiednię czasową tłumienia sygnału znaleziono po dwa instrumenty ze statystycznie istotnie różniącymi się postrzeganą głośnością próbkami.

I tak – w przypadku progu zadziałania, gitara basowa skompresowana z progiem -13,1dBFS okazała się być wyraźnie głośniejsza od skompresowanej z progiem -9,1dBFS., puozn skompresowany z progiem -15,6dBFS okazał się istotnie głośniejszy od puoznu nieskompresowanego. Saksofony zaś – skompresowany z progiem zadziałania -10,8dBFS oraz nieskompresowany okazały się być wyraźnie głośniejsze od saksofonu skompresowanego z progiem zadziałania -12,7dBFS. Wokal, skompresowany z progiem zadziałania -20,8dBFS okazał się być wyraźnie głośniejszy od skompresowanego z progiem zadziałania -16,9dBFS, zaś skrzypce – skompresowane z progiem zadziałania -18dBFS okazały się być głośniejsze od skompresowanych z progiem zadziałania -20,5dBFS. Jak łatwo można zauważyć – w przypadku wokalu i skrzypiec wystąpiło zjawisko przekompresowania.

W przypadku stopnia kompresji, gitara basowa skompresowana ze stopniem 4,2:1 okazała się być istotnie głośniejsza od skompresowanej ze stopniem 2,8:1. Saksofon zaś – skompresowany ze stopniem 3,0:1 okazał się być wyraźnie głośniejszy od saksofonu skompresowanego ze stopniem 1,5:1. Skrzypce – nieskompresowane i skompresowane ze stopniem 3,2:1 okazały się być wyraźnie głośniejsze od skrzypiec skompresowanych ze stopniem 2,1:1 i 1,4:1, zaś wokal, skompresowany ze stopniem 1,9:1 okazał się być głośniejszy od wokalu skompresowanego ze stopniem 1,6:1. Tutaj potwierdziła się sformułowana w założeniach referatu prawidłowość, jednak ciekawe wyniki uzyskane zostały dla skrzypiec – próbka nieskompresowana, jak również najmocniej skompresowna okazały się być głośniejsze od pozostałych.

W przypadku czasu ataku, puozn skompresowany z czzasem ataku 142ms okazał się być głośniejszy od puoznu nieskompresowanego, zaś werbel, skompresowany z czasem ataku 84ms okazał się być wyraźnie głośniejszy od skompresowanego z czasem ataku 32ms. Te - z pozoru absurdalne - wyniki znajdują uzasadnienie w charakterze obwiedni czasowej brzmienia tychże instrumentów w „czystej” postaci.

Dla czasu zwolnienia, puozn nieskompresowany i skompresowany z czasem zwolnienia 82ms, okazał się być wyraźnie głośniejszy od puoznu skompresowanego z czasem zwolnienia 200ms, zaś skrzypce nieskompresowane i skompresowane z czasem zwolnienia 32ms okazały się być istotnie głośniejsze od skrzypiec skompresowanych z czasem zwolnienia 21ms, co by wynikało z pewnego przywyuczajenia ucha ludzkiego do niezniekształconego brzmienia skrzypiec.

4. WNIOSKI I KIERUNKI DALSZYCH DZIAŁAŃ

W ujęciu ogólnym, po przeprowadzonych badaniach możemy sformułować następujące wnioski:

- Z obniżaniem progu zadziałania kompresora, rośnie postrzegana głośność
- Zwiększenie stopnia kompresji powoduje wzrost odbieranej głośności
- Przy instrumentach, o transjentowym charakterze brzmienia, kompresor może stanowić narzędzie kształtowania obwiedni – dłuższy czas ataku powoduje wrażenie głośniejszego nagrania.
- Dobranie zbyt długich czasów zwolnienia powoduje większa bezwładność układu i ujemnie wpływa na odbieraną głośność.
- Wartości parametrów kompresji należy dobrać do charakteru obrabianego materiału!

Nie ulega wątpliwości, że badanie miało charakter zgrubny. By dokładnie wyznaczyć wpływ parametrów kompresji dynamicznej, na percypowaną głośność nagrania instrumentu muzycznego, należałoby przeprowadzić badania dla większej ilości wartości pojedynczego parametru, jak również

dla znacznie poszerzonego instrumentarium. Dla uzyskania naprawdę rzetelnych wyników, takie posunięcie jest konieczne. To z kolei wymagałoby sporządzenia o wiele większej ilości przykładów oraz podzielenia badań odsłuchowych, zakrojonych na większą skalę, na części i sekcje. Niemniej – w dalszej perspektywie – jest to jednak możliwe.

Warto rozważyć też analizę wpływu innych, niż parametry kompresji dynamicznej czynników na postrzeganą głośność nagrania instrumentu muzycznego, jak zawartość zniekształceń harmonicznym w sygnale oraz składowych nieharmonicznych, czy szumu.

LITERATURA

- [1] E. DERUTY - „Dynamic Range' & The Loudness War”,
<https://www.soundonsound.com/sound-advice/dynamic-range-loudness-war> (dostęp 12.04.2019)
- [2] T. P. ZIELIŃSKI, P. KOROHODA, R. RUMIAN - „Cyfrowe przetwarzanie sygnałów w telekomunikacji”,
Wydawnictwo Naukowe PWN SA, Warszawa 2014
- [3] <http://www.studiomastering.net/mastering05.html> (dostęp 28.03.2019 19:32)

EFFECTS OF DYNAMIC COMPRESSION PARAMETERS ON SUBJECTIVE LOUDNESS OF MUSICAL INSTRUMENT RECORDING

There is an increasing trend in music production of decreasing the dynamic range of music recordings. The main purpose is the overall increase of loudness. In most cases, the idea of loudness was referring to the effective value of signal level in a transmission channel. As a result, the material was often treated with aggressive dynamic range compression. The practical experience of audio engineers is, that the subjective impression of loudness is affected not only by the dynamic compression ratio, but by other parameters of compression too. The goal of this work was to investigate, which of the parameters significantly affect the subjective perception of loudness of various music material. It was found that it was possible to set dynamic range compression parameters so that a recording would produce an impression of loudness without excessive dynamic range compression ratio.