# Kompresja stratna audio

## Działanie obu metod

### Metoda μ-law:

* Nieliniowa kwantyzacja – odstęp pomiędzy poziomami nie jest stały, dla niskich i wysokich wartości amplitudy (wartości bliższe -1 i 1) odstęp pomiędzy poziomami jest większy, zmniejsza się precyzja.
* Lepsza jakość niż przy kwantyzacji liniowej (większy zakres dynamiki przy takiej samej liczbie bitów).
* Ograniczony szum kwantyzacji.

### Metoda DPCM z predykcją:

**Do predykcji wykorzystałem medianę z 6 elementów.**

* Zmodyfikowana wersja zwykłej metody DPCM
* Kolejna próbka jest obliczana na podstawie kilku poprzednich elementów, za pomocą wybranej metody, np. średniej arytmetycznej lub mediany (w DPCM jest brana pod uwagę jedynie różnica między kolejnymi próbkami). Kodowane są różnice między kolejnymi próbkami a ich przewidywanymi wartościami.
* Kompresji ulegają różnice między kolejnymi próbkami sygnału dźwiękowego, a nie same próbki.
* Według mnie często słychać większy szum niż w metodzie μ-law ze względu na bardziej nieprzewidywalne zmiany wartości kolejnych próbek (amplituda z większą częstotliwością zmienia się z rosnącej na malejącą i na odwrót).

Mimo tego, że dla większej ilości bitów DPCM wygląda lepiej na wykresach niż μ-law (jest bardziej zbliżony do oryginału) to i tak dźwięk skompresowany metodą μ-law zdaje się brzmieć lepiej.

Przykładowe wykresy dla:

    x = np.linspace(-1, 1, 1000)

    y = 0.9 \* np.sin(np.pi \* x \* 4)

**Kwantyzacja:**

**2 bity**

Obraz zawierający diagram, tekst, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, Wykres, design

Opis wygenerowany automatycznie

**3 bity**

Obraz zawierający diagram, tekst, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, linia, Wykres, Czcionka

Opis wygenerowany automatycznie

**4 bity**

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający diagram, linia, tekst, Wykres

Opis wygenerowany automatycznie

**5 bitów**

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

**6 bitów**

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający linia, tekst, diagram, Wykres

Opis wygenerowany automatycznie

**7 bitów**

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, linia, diagram, Wykres

Opis wygenerowany automatycznie

**8 bitów**

Obraz zawierający tekst, diagram, linia, Wykres

Opis wygenerowany automatycznie

Obraz zawierający tekst, linia, diagram, Wykres

Opis wygenerowany automatycznie

## Wpływ metod kompresji na jakość plików dźwiękowych

### Plik: sing\_low1.wav

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Metoda\Liczba bitów | **8** | **7** | **6** | **5** | **4** | **3** | **2** |
| **μ-law** | Brak większych różnic | Minimalny szum | Trochę mocniejszy szum, nadal słabszy niż dla DPCM, dźwięk stał się głośniejszy | Głośniejszy szum | Podobny szum + dźwięk jest bardziej zniekształcony | Oprócz większego szumu i zniekształcenia dźwięk jest znacznie głośniejszy | Mimo niskiej jakości dźwięk nadal da się zrozumieć |
| **DPCM z predykcją** | Słychać szum | Słychać podobny szum co dla 8 bitów | Mocniejszy szum, nadal treść jest wyraźnie słyszalna | Jeszcze większy szum + dodatkowo trzaski | Pojawiają się trzaski, dźwięk staje się nieprzyjemny dla ucha | Da się rozpoznać tylko sam początek, słychać pisk i trzaski | sam pisk, bardzo głośny, lepiej nie korzystać ze słuchawek |

### Plik: sing\_medium2.wav

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Metoda\Liczba bitów | **8** | **7** | **6** | **5** | **4** | **3** | **2** |
| **μ-law** | Słychać lekki szum | Trochę większy szum | Szum staje się wyraźny | Dalsze wzmocnienie się szumu | Mocniejszy szum, dźwięk trochę zbyt głośny | Mocniejszy szum + dźwięk bardziej zniekształcony | Mocny szum + trzaski. Nadal da się zrozumieć pierwotną treść. |
| **DPCM z predykcją** | Słychać szum choć mniejszy niż dla pliku gorszej jakości, mocniejszy nić dla μ-law | Większy szum | Szum nadal większy niż dla μ-law | Dalsze wzmocnienie się szumu | Dalsze wzmocnienie się szumu | Szum nadal mocniejszy niż dla μ-law, lecz zniekształcenie jest mniejsze | Bardzo głośny dźwięk + bardzo mocny szum. Nadal da się zrozumieć. |

### Plik: sing\_high1.wav

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Metoda\Liczba bitów | **8** | **7** | **6** | **5** | **4** | **3** | **2** |
| **μ-law** | Brak słyszalnej różnicy | Słychać lekki szum w tle | Dźwięk sprawia wrażenie jakby był „poszarpany” | Głośniejszy + zniekształcony dźwięk w niektórych momentach | To samo co dla 5 bitów tylko że mocniej | Bardzo głośny, zniekształcony dźwięk | Nadal słychać że to ten sam krzyk.  Przebijający się dźwięk + trudny do określenia dźwięk w tle |
| **DPCM z predykcją** | lekki szum, znacznie mniejszy niż dla plików gorszej jakości | Minimalnie większy szum niż dla μ-law | W tle słychać ciche „piknięcia” | Głośniejszy + większy szum ale nadal nie jest tak mocny  Jak w poprzednich plikach | To samo co dla 5 bitów + dźwięk trochę zniekształcony | Dźwięk jeszcze głośniejszy niż μ-law ale mniej zniekształcony | Dźwięk jest bardziej stały, brzmi jak równomierny krzyk, mocno różni się pod tym względem od oryginalnego nagrania |

Kompresją metodą **μ-law** prawie zawsze daje lepsze wyniki niż **DPCM z predykcją,** wraz ze wzrostem jakości pliku różnice te się zacierają.

Dźwięk skompresowany za pomocą metody **μ-law** zwykle nadal jest zrozumiały, nawet mimo zniekształcenia i szumu powstałego w wyniku zmniejszenia ilości bitów.

## Jakość dźwięku po działaniu domyślnym (kompresja do 8 bitów)

* Dla metody **μ-law** zwykle nie słychać znacznej różnicy, maksymalnie lekki szum
* Dla metody **DPCM z predykcją** zwykle od razu można usłyszeć co najmniej lekki szum, do jego wychwycenia nie trzeba już przykuwać większej uwagi.