

# 이상 비트율 할당과 신호왜곡 문제점을 고려한 멀티미디어 신호의 연판정 양자화 방법

## Soft-Decision Based Quantization of the Multimedia Signal Considering the Outliers in Rate-Allocation and Distortion

임종욱\*, 노명훈\*, 김무영\*

(Jongwook Lim\*, Myunghoon Noh\*, Moo Young Kim\*)

\*세종대학교 정보통신공학과

(접수일자: 2010년 2월 28일; 수정일자: 2010년 3월 23일; 채택일자: 2010년 4월 10일)

기존 데이터 압축 방식에는 크게 resolution-constrained quantization (RCQ) 방식과 entropy-constrained quantization (ECQ) 방식이 있다. RCQ 방식은 고정 비트율 전송을 가능하게 하지만 셀 크기의 변화에 따른 이상 신호왜곡이 발생하며, ECQ 방식은 셀 크기가 고정된 대신에 이상 비트율 할당 문제가 발생한다. 본 논문에서는 기존 RCQ 방식의 대표적인 학습기법인 generalized Lloyd algorithm (GLA)을 개선한 cell-size constrained vector quantization (CCVQ) 방식을 제안한다. CCVQ 알고리즘은 셀 크기에 따라 유동적으로 페널티 척도를 주는 방식으로 기존의 RCQ와 ECQ 사이의 soft-decision을 가능하게 한다. 제안 알고리즘을 사용할 경우 기존의 GLA에 비해 약간의 평균왜곡 증가는 발생하나 이상 신호왜곡을 줄일 수 있다.

**핵심용어:** High-Rate 이론, 소스 코딩, 양자화, GLA, 이상 신호왜곡

**투고분야:** 음성처리 분야 (2,1)

There are two major conventional quantization algorithms: resolution-constrained quantization (RCQ) and entropy-constrained quantization (ECQ). Although RCQ works well for fixed transmission-rate, it produces the distortion outliers since the cell sizes are different. Compared with RCQ, ECQ has the constraints on the cell size but it produces the rate outliers. We propose the cell-size constrained vector quantization (CCVQ) that improves the generalized Lloyd algorithm (GLA). The CCVQ algorithm is able to make a soft-decision between RCQ and ECQ by using the flexible penalty measure according to the cell size. Although the proposed method increases the small amount of overall mean-distortion, it can reduce the distortion outliers.

**Keywords:** High-Rate Theory, Source Coding, Quantization, GLA, Distortion Outliers

**ASK subject classification:** Speech Signal Processing (2,1)

### I. 서론

데이터를 압축 및 전송하는데 사용하는 양자화 방식은 고정 비트율 (fixed-rate)을 사용하는 resolution-constrained quantization (RCQ) 방식과 가변 비트율 (variable-rate)을 사용하는 entropy-constrained quantization (ECQ) 방식이 있다 [1-5]. RCQ는 전송하는 각 심볼의 비트율을 고정시켜 평균 왜곡을 최소화 시키는

양자화 방식이고, 이에 각 보로노이 영역 (Voronoi region)의 크기는 데이터 포인트의 분포에 따라 서로 다른 크기를 가지게 된다. 반면 가변 비트율을 사용하는 ECQ는 전체 평균 왜곡을 제한하여 평균 왜곡을 최소화하는 방식이고 이상적인 ECQ 방식은 lattice quantizer와 같은 uniform quantization과 허프만 코딩과 같은 무손실 압축 방법의 결합이라고 알려져 있다. 따라서 ECQ 방식을 사용할 경우 각 보로노이 영역은 크기가 일정하게 되고 데이터 포인트의 분포가 높은 심볼에 적은 비트수를 할당하여 양자화하게 된다.

위와 같은 특성에 의해 기존 네트워크 환경인 써킷망

책임저자: 김 무 영 (mooyoung@sejong.ac.kr)  
143-747 서울서 광진구 군자동 세종대학교 정보통신공학과  
(전화: 02-3408-3695; 팩스: 02-3408-4330)

(circuit-switched communication channels)에서는 고정된 비트율 양자화하는 RCQ 방식의 사용이 효율적이었으나, 최근 사용량이 증가하고 있는 패킷망 (packet-based communication channels)에서는 평균 비트율을 고정시켜 평균 왜곡을 최소화 하는 ECQ 방식이 보다 효율적인 양자화 방식으로 쓰이게 되었다.

반면, RCQ와 ECQ 방식은 모두 데이터 포인트의 발생 확률이 높은 부분에 코딩 효율을 높이고, 발생확률이 낮은 부분에 코딩 효율을 낮춤으로써 전체적으로는 평균 코딩 효율을 높이고자 하는 공통점이 있다. 이와 같은 불균형한 코딩 방식은 각각의 경우 다음과 같은 문제점을 발생시킨다. RCQ의 경우에는 발생확률이 낮은 영역인 보로노이 영역의 크기가 큰 부분에서 이상 신호왜곡 (outliers in distortion)을 초래하게 되고, ECQ의 경우에는 발생확률이 낮은 보로노이 셀에서 이상 비트율 할당 (outliers in rate allocation)을 초래하게 된다. RCQ에서 나타나게 되는 이상 신호왜곡은 평균왜곡 (mean distortion)과 상반되는 개념으로 비록 그 발생확률은 낮지만 인간의 심리청각 특성상 해당 시스템의 심각한 성능저하를 야기하게 되며 [6], ECQ에서 나타나는 이상 비트율 할당은 평균 비트율 (average rate)과 상반되는 개념으로써 패킷망의 congestion을 발생시켜 심각한 네트워크 효율의 저하를 야기하게 된다 [7]. 반면에, RCQ는 이상 비트율 할당을 발생시키지 않고, ECQ는 이상 신호왜곡을 발생시키지 않게 된다.

기존에는 시스템 디자이너가 신호왜곡의 증가나 비트율 증가 중 중요 척도에 따라 RCQ와 ECQ 가운데 한가지만을 hard-decision 방식으로 선택해서 사용해야만 했다. 이에 이상 신호왜곡과 이상 비트율 할당을 조절할 수 있는 유동적인 페널티 척도 (flexible penalty measure)를 도입하여 RCQ와 ECQ 사이의 새로운 척도를 찾고자 하는 연구들이 진행되어 왔다. 그 중 sensitivity matrix를 이용하여 weighted mean-squared error (MSE) 척도를 제안한 방법이 제안되었으나 이상 신호왜곡에 대한 사항이 고려되지 않았으며 [8], just noticeable difference (JND) 개념을 도입하여 평균왜곡과 이상 신호왜곡의 관점에서 개선을 이룬 연구도 있었으나 [9], 구현 알고리즘이 오직 1차원인 scalar quantization (SQ)에만 적용되는 약점을 가졌다.

이에 본 논문에서는 셀 크기에 따라 제한을 주는 cell-size constrained vector quantization (CCVQ)방식을 제안한다. 일반적인 GLA는 각 셀의 내부에서 발생되는 데이터 포인트의 확률에 따라 보로노이 영역을 형성하

게 되는데, CCVQ는 크기가 작은 셀에 대해 penalty를 주어 최적의 코드북을 찾으려 하였다. 이에 따라 이상 신호 왜곡을 줄일 수 있다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 기본적인 high-rate 이론에 대해서 설명하고 3장에서는 본 논문의 바탕이 되는 JND 기반 RCQ 방식에 대한 설명이 있겠다. 다음 4장에서는 제안하는 알고리즘에 대한 자세한 설명을 하고, 5장에서 기존의 GLA와 제안된 알고리즘에 대한 비교 실험 결과를 제시하고, 마지막 6장에서 최종 결론을 짓도록 하겠다.

## II. 선행 연구

본 절에는 기본적인 high-rate 이론과 제안 알고리즘의 동기가 되는 JND 기반의 RCQ에 대한 설명이 있겠다.

먼저 High-rate 이론을 설명하는데 있어서 다음을 가정한다. 입력신호  $x^k$ 는 probability density function (PDF)이  $f(x^k)$ 인 랜덤벡터를 의미하며  $k$ 는 입력신호의 차원을 의미한다. 양자화 과정 중 인코딩 과정에 해당하는  $Q$ 는  $k$  차원의 입력신호  $x^k$ 를  $N$ 개의 centroids 중 한 개로 맵핑시키는 함수이다. 이런 경우  $i$  번째 centroid  $c_i^k$ 의 보로노이 영역 (Voronoi region)  $V_i$  는

$$V_i = \{x^k \in R^k : Q(x^k) = c_i^k\} \quad (1)$$

와 같이 표현할 수 있다.

본 논문에서 사용하는  $r$  차승 평균 왜곡척도 (mean  $r$ -th power distortion measure)는 다음의 식과 같이 표현된다:

$$d(x^k, Q(x^k)) = \|x^k - Q(x^k)\|_r = \left( \frac{1}{k} \sum_1^k (x^k - Q(x^k))^2 \right)^{r/2}. \quad (2)$$

주어진 왜곡척도에 대해서 최적의 보로노이 영역은 다음과 같이 표현할 수 있다:

$$V_i = \{x^k \in R^k : \|x^k - c_i^k\|_r \leq \|x^k - c_j^k\|_r, \text{ for all } j\}. \quad (3)$$

High-rate 이론은 양자화 비트수가 충분하며, 각 보로노이 영역 내의 데이터 PDF는 일정한 값을 가짐을 가정한다.

입력신호  $x^k$ 가  $i$  번째 보로노이 영역 내에 포함되어 있을 확률은 다음과 같이 정의할 수 있다:

$$P_i = \int_{V_i} f(x^k) dx^k \approx f_i v(V_i). \tag{4}$$

따라서,  $i$  번째 셀의 평균 왜곡은

$$\begin{aligned} D_i &= \frac{\int_{V_i} f_{X^k}(x^k) \|x^k - \hat{x}\|_r dx^k}{\int_{x^k \in R^k} f_{X^k}(x^k) dx^k} \\ &\approx V_i^{-1} \int_{V_i} \|x^k - c_i^k\|_r dx^k \\ &\approx C(r, k, V_i) v(V_i)^{r/k} \end{aligned} \tag{5}$$

로 표현 가능하며, 여기서  $v(V_i)$ 와  $C(r, k, V_i)$ 는 각각  $i$  번째 보로노이 영역의 체적과 양자화 계수를 의미한다. 또한  $V_i$ 는 Gersho 계수로  $V_i = V$ 로 가정할 수 있으며, 이 계수는 보로노이 영역의 기하학적 특성과 관련이 있다 [1].

따라서 이것을 바탕으로, 모든 셀에 대한 평균 왜곡을 계산하면 다음과 같다.

$$\begin{aligned} D &= \sum_{i=1}^I P_i D_i \\ &\approx C(r, k, V) \int_{x^k \in R^k} f_{X^k}(x^k) g_r(x^k)^{r/k} dx^k, \end{aligned} \tag{6}$$

여기서  $g_r(x^k)$ 는 centroid density이다.

RCQ는 고정된 비트율을 사용하기 때문에 centroid의 총 숫자가  $N$ 으로 고정되어 있으므로 다음과 같은 제약이 가진다:

$$\int_{x^k \in R^k} g_r(x^k) dx^k = N. \tag{7}$$

위 식을 (6)에 대입하여 오일러-라그랑제 방정식 (Euler-Lagrange equation)을 풀면

$$g_r(x^k) = N \frac{f_{X^k}(x^k)^{\frac{k}{k+r}}}{\int_{x^k \in R^k} f_{X^k}(x^k)^{\frac{k}{k+r}} dx^k} \tag{8}$$

와 같은 최적 centroid density를 구할 수 있다. (8)에서 구한 값을 (6)에 대입하여 최적의 평균 왜곡을 구하면

$$D_{RC} = C(r, k, V) N^{-\frac{r}{k}} \left( \int_{x^k \in R^k} f_{X^k}(x^k)^{\frac{k}{k+r}} dx^k \right)^{\frac{k+r}{k}}$$

$$\approx C(r, k, V) e^{\frac{r}{k} R} \left( \int_{x^k \in R^k} f_{X^k}(x^k)^{\frac{k}{k+r}} dx^k \right)^{\frac{k+r}{k}} \tag{9}$$

와 같이 계산되며, 여기서  $R$ 은 양자화 시 필요한 비트 수를 나타낸다.

다음으로는 기존에 연구된 방식 중 이상 신호왜곡 관점에서 개선을 이룬 JND 기반 RCQ 방식에 대해서 살펴볼 도록 하겠다 [9]. JND란 인간의 시각이나 청각 등의 감각 기관에서 정의된 개념으로 인간의 감각 기관이 아주 작은 왜곡은 인식하지 못하고 어떤 문턱값을 넘어선 순간부터 왜곡 증가에 따라 왜곡을 인식하게 되는 것을 말한다. 그림 1에서 일반적인 squared error와 JND 기반의 squared error의 차이를 확인할 수 있다. 그림에서 확인할 수 있듯이 어떤 문턱값까지는 JND 기반의 squared error가 0의 값을 가지다가 어느 값 이상이 되면 인식되는 error의 값이 증가함을 확인할 수 있다. [9]에서 JND 기반 RCQ는 다음 수식으로 정의된다.

$$d_p(x^k, Q(x^k)) = \begin{cases} 0 & , \text{if } \|x^k - Q(x^k)\|_r < d_{JND} \\ \|x^k - Q(x^k)\|_r - d_{JND} & , \text{elsewhere} \end{cases} \tag{10}$$

$\|x^k - Q(x^k)\|_r < d_{JND}$ 에 포함되는  $c_i$  주변의 영역을  $T_i$ 라고 정의하자. 따라서  $V_i$ 의 영역 안에 있는 점들 중에  $T_i$ 에 포함되지 않은 영역은  $V_i \cap T_i^c$ 로 나타낼 수 있다. 이것을 바탕으로  $c_i^k$  주변 양자화 셀  $V_i$ 의 JND 기반 RCQ 평균왜곡은 다음의 식으로 정리된다:

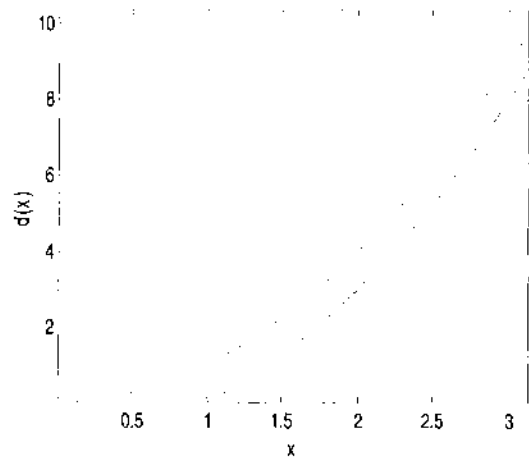


그림 1. squared error (dashed line)와 JND 기반의 error (solid line)의 예시

Fig. 1. Examples of squared error (dashed line) and JND based error (solid line).

$$\begin{aligned}
 D_j(i) &\approx \frac{1}{v(V_i)} \int_{V_i} d_p(x^k, c_i^k) dx^k \\
 &= C(k, r, V_i) v(V_i)^{r/k} - C(k, r, V_i \cap T_j) \frac{v(V_i \cap T_j)^{r/k+1}}{v(V_i)} \\
 &\quad - d_{JND} \frac{v(V_i \cap T_j)}{v(V_i)}. \tag{11}
 \end{aligned}$$

이 때, 모든 셀에 대한 평균왜곡은 다음과 같이 나타낼 수 있다:

$$\begin{aligned}
 D &= \sum_{i, j \in T} P_j(i) D_j(i) \\
 &\approx C(k, r, V) \int_{g_c(x^k) \leq 1/v(T)} f_{X^k}(x^k) g_c(x^k)^{-r/k} dx^k \\
 &\quad - \bar{C}(k, r, T) \int_{g_c(x^k) < 1/v(T)} f_{X^k}(x^k) g_c(x^k) dx^k \\
 &\quad - d_{JND} \int_{g_c(x^k) < 1/v(T)} f_{X^k}(x^k) dx^k. \tag{12}
 \end{aligned}$$

위 식에서  $\bar{C}(k, r, T) = C(k, r, T) v(T)^{r/k+1} - d_{JND} v(T)$  이며, centroid density는  $g_c(x^k) \approx 1/v(V_i)$ ,  $x^k \in V_i$ 의 값을 가진다.

위 (12)를  $N$ 개의 양자화 점으로 최소화하는 이상적인 centroid density는 오일러-라그랑제 방정식을 이용하여 계산하면 다음과 같다:

$$\begin{aligned}
 \eta &= \int_{x^k \in R^k} [C(k, r, V) f_{X^k}(x^k) g_c(x^k)^{-\frac{r}{k}} \\
 &\quad - \bar{C}(k, r, T) f_{X^k}(x^k) g_c(x^k) - d_{JND} f_{X^k}(x^k) \\
 &\quad + \lambda g_c(x^k) + u(x^k) g_c(x^k)] dx^k \tag{13}
 \end{aligned}$$

이 때,  $u(x^k)$ 는 다음과 같다:

$$u(x^k) \begin{cases} = 0, & \text{for } g_c(x^k) < \frac{1}{v(T)} \\ > 0, & \text{for } g_c(x^k) = \frac{1}{v(T)} \end{cases} \tag{14}$$

(13)을 오일러-라그랑제 방정식을 이용하여 계산하면 다음과 같은 이상적인 centroid density를 계산할 수 있다:

$$\begin{aligned}
 \lambda \cdot \frac{r}{k} C(k, r, V) f_{X^k}(x^k) g_c(x^k)^{-(k+r)/k} - \bar{C}(k, r, T) f_{X^k}(x^k) \\
 \begin{cases} = 0, & \text{for } g_c(x^k) < \frac{1}{v(T)} \\ < 0, & \text{for } g_c(x^k) = \frac{1}{v(T)} \end{cases} \tag{15}
 \end{aligned}$$

따라서,

$$g_c(x^k) = \begin{cases} \left( \frac{\frac{r}{k} C(k, r, V) f_{X^k}(x^k)}{\lambda - \bar{C}(k, r, T) f_{X^k}(x^k)} \right)^{\frac{k}{k+r}}, & \text{if } g_c(x^k) < \frac{1}{v(T)} \\ \frac{1}{v(T)}, & \text{elsewhere} \end{cases} \tag{16}$$

와 같은 최종 centroid density를 구할 수 있다.

### III. 제안 알고리즘

앞서 설명한 high-rate 이론과 기존에 연구되었던 JND 기반 RCQ를 바탕으로 본 절에서는 개선된 알고리즘을 제안한다. 제안 알고리즘은 RCQ 방식을 사용하는 대표적인 학습기법인 GLA를 개선하여 이상 신호왜곡을 줄이는 cell-size constrained vector quantization (CCVQ) 방식이다.

일반적인 GLA는 입력신호  $x^k$ 와 코드벡터  $c^k$ 로 다음과 같은 왜곡 척도를 이용하여 최적의 코드북을 생성하게 된다:

$$\eta = \|x^k - c^k\|_r. \tag{17}$$

(17)으로 보로노이 영역을 구하는 GLA 알고리즘은 데이터 포인트의 발생확률에 따라 데이터 포인트가 많이 발생하는 영역에서는 크기가 작은 셀을 만들게 되고, 데이터 포인트의 발생확률이 낮은 영역에서는 크기가 큰 셀을 만들게 된다. 이런 GLA의 불균일한 코딩 효율의 분배는, 상대적으로 발생확률이 낮은 데이터 포인트에 대해서 이상 신호왜곡을 초래하게 되므로, 인간의 심리청각 특성상 해당 시스템의 심각한 성능저하로 평가되어진다.

그러나, 앞 절에서 확인한 JND 기반 RCQ 방식에 따르면 JND 특성에 의해 어느 정도의 작은 왜곡은 인간의 청각 특성에 의해 감수될 수 있다. 즉, 크기가 어느 정도 작은 셀에 대해서 추가적인 왜곡이 가해져도 그 왜곡은 인간의 심리청각 특성상 큰 차이를 느끼지 못하게 된다. 따라서, JND의 특성을 이용하여 각 보로노이 영역, 즉 셀 크기에 penalty를 가하면 이상 신호왜곡을 줄일 수 있게 된다. 크기가 어느 정도 작은 셀에 인위적으로 왜곡을 추가하여 입력 데이터의 발생 확률이 높은 부분에서 코드 벡터를 결정할 때 강제로 최소 왜곡의 코드 벡터를 선택하는 대신, 외곽의 코드 벡터를 선택하게 함으로써 입력 벡터를 외곽으로 퍼지게 강요하게 된다. 이러한 penalty

는 결국 외곽에서도 코드벡터를 만들게 되어 이상 신호외곽을 줄이는 효과를 기대할 수 있다. 이런 이상 신호외곽의 감소는 최종적으로 distortion outlier를 줄이는 효과를 가져 오게 된다.

이상 신호외곽을 줄이는 CCVQ 알고리즘은 다음과 같은 왜곡척도를 이용해서 코드벡터를 구한다:

$$\eta = \|x^k - c_i^l\|_r + \lambda f(D_i) \quad (18)$$

여기서  $\lambda$ 는 penalty를 가하는 정도를 조절하는 변수로써 이상 신호외곽의 감소 정도를 결정하게 된다. 앞서 설명했듯이 크기가 작은 셀에 대해서는 어느 정도의 왜곡을 강제로 추가함으로써 코드벡터를 결정하는데, 있어서 penalty를 부과하게 된다. 즉 셀에 크기에 따라 비례해서 왜곡을 가하는 것이 아니라 어느 정도 작은 부분에는 강제로 왜곡을 추가해서 입력 벡터를 외곽으로 퍼지게 한다. 또한 penalty를 가하는 평균왜곡 함수는 그림 2와 같은 왜곡함수  $f(\cdot)$ 를 사용하였다. 본 실험에서  $f(D_i)$  함수의 0이 되는 지점은 실험적으로 구한 값이며, distortion outlier의 정의나 입력 데이터의 종류가 바뀔 경우에는 추가적인 실험으로 변경된 값을 사용하도록 하여야 한다. 이처럼 제안 알고리즘은 평균왜곡의 증가를 발생시키기는 하지만, 데이터 포인트의 발생확률이 낮은 외곽 부분에서도 조밀하게 셀을 만들게 되어 이상외곽 측면에서 이점을 가지게 된다. 이에 따라 CCVQ의 보로노이 영역은 다음과 같이 결정하게 된다:

$$V_i^{(t)} = \left\{ \begin{array}{l} \|x^k - c_i^{(t-1)}\|_r + \lambda f(D_i^{(t-1)}) \\ \leq \|x^k - c_m^{(t-1)}\|_r + \lambda f(D_m^{(t-1)}), \forall m, i \in I, m \neq i \end{array} \right\} \quad (19)$$

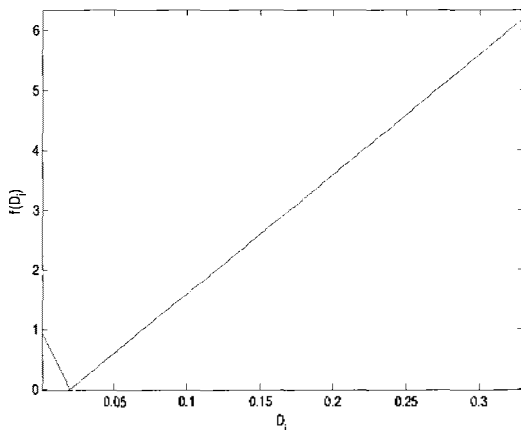


그림 2. 가중치 함수  $f(D_i)$   
Fig. 2. Weighting function  $f(D_i)$ .

제안 알고리즘은 표 1에 정리되어 있다. 우선 iteration 변수인  $t$ , 셀별 최대왜곡  $D_i^{(0)}$ , 전체 평균왜곡  $D^{(0)}$ , 그리고 코드북  $\{c_i^{(0)}, i = 1, 2, \dots, I\}$ 을 초기화하게 된다. 여기서  $I$ 는 코드벡터의 개수로써 양자화시 사용하는 비트에 의해 결정된다. 또한 제안 알고리즘인 CCVQ에서 사용하는 penalty 함수  $f(\cdot)$ 를 정의한다. 다음으로는 앞서 정의된 초기값들을 이용하여 현 단계의 보로노이 영역  $V_i^{(t)}$ 을 결정하게 된다. 그것을 바탕으로 코드벡터  $c_i^{(t)}$ 를 구하게 되고, 구한 코드벡터로 전체 평균왜곡과 셀별 최대왜곡을 측정하게 된다. 최종적으로  $(t-1)$  단계의 전체 평균왜곡과  $(t)$  단계의 전체 평균왜곡을  $(f)$ 와 같이 비교하고 특정 문턱값 이하의 값을 만족시킬 때까지 위의 과정을 반복하여 최종적으로 이상적인 코드벡터를 생성하게 된다. 본 논문에서  $\epsilon$ 은 0.0005를 사용하였다.

### IV. 실험 및 결과

본 절에서는 제안 알고리즘을 구현하여 실험 결과를 확인해 본다. 실험에 사용하는 입력 데이터로는  $N(0,1)$  분포를 가지는 2차원 single Gaussian 데이터를 사용하였다.

첫 번째로 일반적인 GLA와 제안 알고리즘의 centroid 점들의 분포와 그에 따른 보로노이 영역을 확인하였다.

표 1. CCVQ 알고리즘  
Table 1. CCVQ algorithm.

(a) 초기화 단계
- $t=1, D^{(0)}=0, D_i^{(0)}=1$
- 코드북 $\{c_i^{(0)}, i = 1, 2, \dots, I\}$ 초기화
- $f(x) = \begin{cases} 1-50x & , x < 0.02 \\ 20x - 0.02 & , otherwise \end{cases}$
(b) 보로노이 영역 결정
$V_i^{(t)} = \left\{ \begin{array}{l} \ x^k - c_i^{(t-1)}\ _r + \lambda f(D_i^{(t-1)}) \\ \leq \ x^k - c_m^{(t-1)}\ _r + \lambda f(D_m^{(t-1)}), \forall m, i \in I, m \neq i \end{array} \right\}$
(c) 코드벡터 생성
$c_i^{(t)} = \operatorname{argmin}_{y^k \in R^L} \sum_{x^k \in V_i^{(t)}} (\ x^k - y^k\ _r + \lambda f(D_i^{(t-1)}))$
(d) 전체 평균왜곡 측정
$D^{(t)} = \sum_{i=1}^I \sum_{x^k \in V_i^{(t)}} (\ x^k - c_i^{(t)}\ _r + \lambda f(D_i^{(t-1)}))$
(e) 셀별 최대왜곡 측정
$D_i^{(t)} = \max_{x^k \in V_i^{(t)}} \ x^k - c_i^{(t)}\ _r$
(f) if $ (D^{(t-1)} - D^{(t)})/D^{(t)}  < \epsilon$ , quit. Otherwise, set $t=t+1$ and go to (b).

그림 3은 2차원의 single Gaussian 데이터를 8비트로 양자화 할 때, 일반적인 GLA와 CCVQ가 갖는 centroid와 보로노이 영역의 형태를 보여주고 있다. 제안 알고리즘은 기존 GLA에 비해 centroid들이 중심에서부터 밖으로 퍼지는 효과를 나타내며 그에 따라 안 쪽의 보로노이 영역의 크기가 커지는 것을 확인할 수 있다. 이에 따라 기존 GLA에 비해 외곽에도 보로노이 영역을 가짐으로써 이상 신호왜곡을 줄이게 된다. 본 실험에서  $\lambda$ 는 0.0129를 사용하였다.

그림 4는 평균이 각각  $(-0.5, -0.5)$ ,  $(0.5, 0.5)$ 에 위치하는 두 개의 mixed Gaussian 데이터에 대한 보로노이 영역을 보여주고 있다. GLA와 CCVQ 모두 해당 데이터에 대해서도 정상적으로 동작함을 확인할 수 있다. 하지만, CCVQ는 일반적인 VQ training 알고리즘인 GLA와 마찬가지로 local minimum에 수렴하는 한계를 가지는 알고리즘이므로 입력 데이터가 여러 개의 중심을 가지면서 완전히 이산적으로 분포하는 경우에는 최적의 코드북을 찾기가 어렵다. CCVQ와 GLA의 이러한 문제점은 다양한 초기 벡터를 이용하여 다수의 local minima를 찾고 이 중에 최적의 해를 global minimum으로 선택하는 방법 등을 적용한다면 부분적으로 해결할 수 있을 것이다.

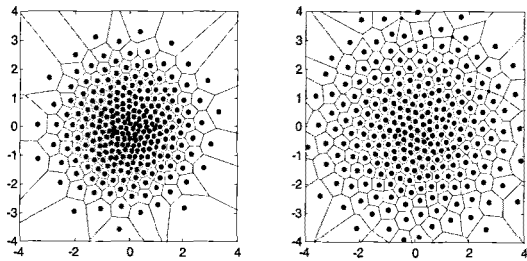


그림 3. Single Gaussian에 대한 GLA와 CCVQ의 Voronoi regions  
Fig. 3. Voronoi regions of GLA and CCVQ for single Gaussian data.

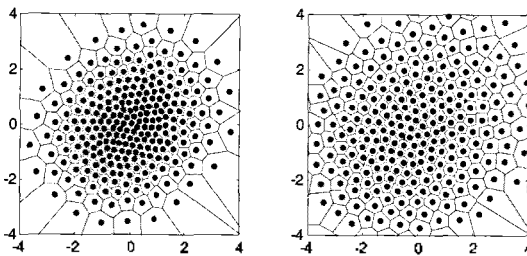


그림 4. Mixed Gaussian에 대한 GLA와 CCVQ의 Voronoi regions  
Fig. 4. Voronoi regions of GLA and CCVQ for mixed Gaussian data.

다음으로 제안 알고리즘 효과를 분석하기 위해 GLA와 제안 알고리즘 각각의 왜곡 분포를 비교하여 보았다. 그림 5에 GLA와 제안 알고리즘에서 나타나는 왜곡의 PDF와 cumulative mass function (CDF)이 나타나 있다. 그림에서와 같이 기존 GLA에 비해 제안 알고리즘의 왜곡 PDF가 작은 영역에서는 적은 분포를 갖다가 큰 영역으로 갈수록 비중이 커지는 것을 확인할 수 있다. 이것은 제안 알고리즘이 GLA에 비해 보로노이 영역이 큰 부분의 분포가 높다는 것을 의미하며, 즉 CDF 값이 커지는 지점에서부터 제안 알고리즘이 기존 GLA에 비해 이상 신호왜곡을 줄이는 것으로 해석할 수 있다. 본 실험에서  $\lambda$ 는 0.0129를 사용하였다.

다음은 기존 GLA와 제안 알고리즘의 성능 비교이다. 성능 비교 척도로는 각 알고리즘의 평균왜곡과 이상 신호 왜곡이 전체 왜곡에서 차지하는 비율을 구한 값을 사용하였다. 표 2에서 성능의 비교 결과를 확인할 수 있으며, 이 실험에서  $\lambda$ 는 0.0129를 사용하였다. 제시된 결과와 같이 평균 왜곡의 관점에서는 제안 알고리즘이 기존 GLA에 비해 0.002 정도 증가하는 것을 확인할 수 있다. 그러나 전체왜곡 중 이상 신호왜곡을 제외한 왜곡의 분포가 기존 GLA에 비해 제안 알고리즘이 높은 것을 확인할 수

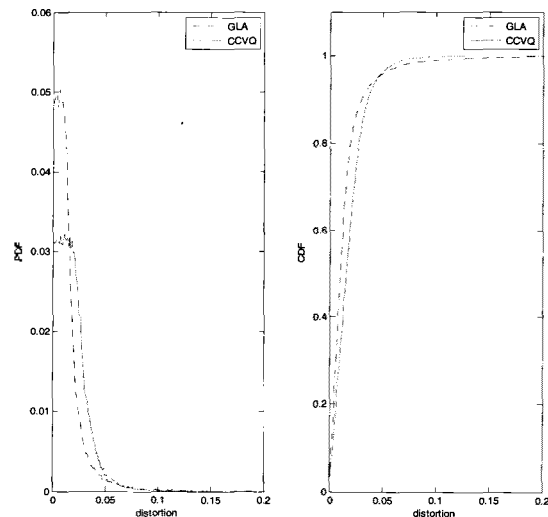


그림 5. GLA와 CCVQ의 왜곡 PDF와 CDF  
Fig. 5. Distortion PDF and CDF of GLA and CCVQ.

표 2. GLA와 CCVQ의 성능 비교  
Table 2. Experimental performance of GLA and CCVQ.

	평균 왜곡	전체 왜곡 중 이상 신호왜곡을 제외한 비율 (%)	
		< 0.075	< 0.150
GLA	0.016	98.02	99.47
CCVQ	0.018	99.21	99.96

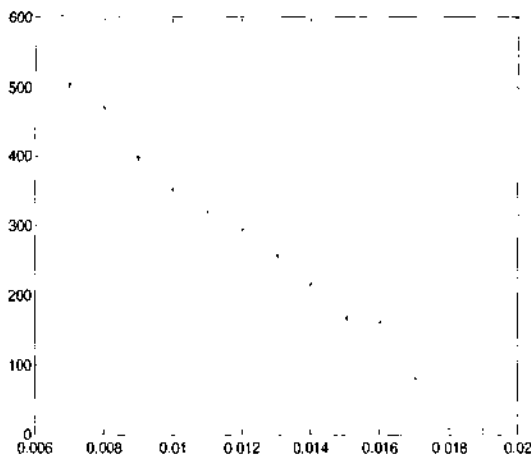


그림 6.  $\lambda$ 에 따른 코드북 크기의 변화  
Fig. 6. Codebook size variation as  $\lambda$  changes.

있으며, 이는 제안 알고리즘이 GLA에 비해 이상 신호왜곡을 줄이는 것으로 해석 할 수 있다.

그림 6에서는  $\lambda$ 에 따른 코드북의 크기 변화를 확인 할 수 있다.  $\lambda$ 가 증가함에 따라 코드북의 크기가 대체로 감소하는 것을 확인 할 수 있으며, 이는 즉  $\lambda$ 가 커질수록 penalty가 증가하기 때문이다. 이에 따라 크기가 작은 보로노이 영역들이 더 많이 소멸되어 평균 비트율이 점점 떨어지는 것으로 해석 할 수 있다. 본 실험에서는 제안 알고리즘을 8비트 GLA와 비교하고자 하였다. 제안 알고리즘의 코드북 크기는  $\lambda$  값에 따라 결정되므로, 공정한 비교를 위하여 8비트 코드북이 생성되도록  $\lambda$  값을 0.0129로 선택하였다.

결론적으로 제안하는 CCVQ 알고리즘은 기존 GLA에 비해 약간의 평균왜곡 증가는 발생시키나 이상 신호왜곡을 줄이는 효과를 얻게 되며, 이는 향후 perceptual quality의 향상을 기대 할 수 있다. 또한 high-rate 이론에 의하면 본 논문에서 실험하지 않은 이상 비트율 할당의 경우도 신호왜곡의 결과와 유사함을 예상할 수 있고, 따라서 이상 신호왜곡을 줄이듯 이상 비트율 할당을 줄일 수 있음을 기대할 수 있다.

## V. 결론

본 논문에서는 이상 신호왜곡을 줄이기 위하여 셀 크기에 따라 제한을 두는 CCVQ를 제안한다. 제안하는 알고리즘은 기존에 사용하던 RCQ와 ECQ 중 하나의 방식을 선택하는 Hard-Decision 방식이 아닌 두 방식의 중간 개념인 Soft-Decision 방식이다. 또한 본 논문에서 제안하는

알고리즘은 평균왜곡의 관점에서는 어느 정도의 손해를 감수하지만, 이상 신호왜곡의 관점에서는 이득을 얻는 효과를 가지게 된다. 제안된 CCVQ 방식은 근본적인 양자화 방식의 개선이기 때문에 향후 다양한 응용처에서 사용이 가능하며, 또한 음성, 영상 그리고 오디오 등 다양한 멀티미디어 데이터의 압축 및 전송 등에 사용될 수 있다. 향후 CCVQ 알고리즘을 실제 음성 및 오디오 코딩에 적용하여 인간의 청각 특성에 따른 perceptual quality 관점의 실험을 진행할 예정이다.

## 감사의 글

이 논문은 2008년 정부(교육과학기술부)의 재원으로 학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2008-331-D00430).

## 참고문헌

1. A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 4, pp. 373-380, 1979.
2. P. L. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 139-149, 1982.
3. T. D. Lookabaugh and R. M. Gray, "High-resolution quantization theory and the vector quantizer advantage," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 1020-1033, 1989.
4. R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325-2383, 1998.
5. W. B. Kleijn, *A Basis for Source Coding: Course Notes*, KTH, Stockholm, 2008.
6. V. Grancharov and W. B. Kleijn, "Speech quality assessment," *Handbook of Speech Processing*, J. Benesty, A. Huang, and M. Sondhi, Eds, Springer, ch. 5, pp. 83-102, 2007.
7. S. Kweon, K. Shin, and Z. Zheng, "Statistical real-time communication over ethernet for manufacturing automation systems," in Proc. Fifth *IEEE Real-Time Techn. Apps. Symp.*, Vancouver, pp. 192-202, 1999.
8. W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 5, pp. 367-381, 1995.
9. M. Y. Kim, and W. B. Kleijn, "Resolution-Constrained Quantization With JND-Based Perceptual-Distortion Measures," *IEEE Signal Process. Letters*, vol. 13, no. 11, pp. 703-706, 2006.

---

## 저자 약력

---

• **임 종 욱 (Jongwook Lim)**

2001.3-2008.8 세종대학교 정보통신공학과, 학사  
2008.8~현재 세종대학교 정보통신공학과, 석사과정



• **노 명 훈 (Myunghoon Noh)**

2003.3-2010.2 세종대학교 정보통신공학과, 학사  
2010.2~현재 세종대학교 정보통신공학과, 석사과정



• **김 무 영 (Moo Young Kim)**

1989.3-1993.2 연세대학교 전자공학과, 학사  
1993.3-1995.2 연세대학교 전자공학과, 석사  
1995.2-2000.12 삼성종합기술원 전문연구원  
2001.1-2004.11 Royal Institute of Technology (KTH, 스웨덴) Dept.  
Signals, Sensors, Systems, 박사  
2004.12-2005.2 Royal Institute of Technology (KTH, 스웨덴) Dept.  
Signals, Sensors, Systems, PostDoc  
2005.2-2006.8 Ericsson Research (스웨덴), Senior Research Engineer  
2006.8-현재 세종대학교 정보통신공학과, 조교수  
※ 관심분야: 음성/오디오/비디오 신호처리, 패턴인식, 정보이론.