개선된 부호책 갱신 방법을 이용한 VQ 학습 알고리즘

VQ Design Algorithm Using Modified Codebook Updating Method

백 성 준*, 최 용 진*, 이 주 헌*, 성 굉 모*
(Seon Joon Baek*, Yong jin Choi*, Joo Hun Lee*, Koeng Mo Sung*)

*본 연구는 현대전자(주)의 지원으로 이루어졌습니다.

요 약

본 논문에서는 기존에 제시된 수정된 K-평균 방법을 이용한 VQ 학습 알고리즘을 분석하고, 보다 개선된 성능을 보이는 학습 알고리즘을 제안한다. 수정된 K-평균 학습 알고리즘은 자기 집단에 속하는 데이터의 중심을 새로운 코드워드로 삼는 것이 아니라 현재 코드워드와 새로 구한 집단의 중심을 연결한 선상에서 새로 구한 중심 너머의 일정한 점을 새로운 코드워드로 선택하는 방식이다. 본 논문에서는 이렇게 구한 새로운 코드워드가 어떠한 조건을 만족할 때 알고리즘이 반복적 감소의 성질을 가지는지 살펴보고, 그 조건을 만족시키는 영역 중 기존의 방식보다 더 좋은 성능을 보이는 코드워드 선택법을 제시함으로써 개선된 학습 알고리즘을 제안한다.

ABSTRACT

In this paper, we investigate the modified K-means method(MKM) for VQ training algorithm and present a new training algorithm. The MKM algorithm do not select current centroid as a new codeword. Instead, it select a certain point on the line from current codeword to the centroid as a new codeword. In this work, the condition for iterative descent property of MKM algorithm is analyzed and an improved algorithm with differently selected codeword is presented.

I. 서 론

벡터양자화 기법은 음성압축이나 영상압축, 음성인식, 화자인식, 패턴인식 등에 널리 쓰는 알고리즘 중 하나다[1]. 벡터양자화 기법의 장점 중 하나는 이에 쓰이는 복호화기 및 부호화기의 구조가 매우 간단하다는 것이다. 부호화 하고자 하는 데이터는 먼저 벡터열의 집합으로 구성된다. 그 후 이 벡터 열을 이용하여 반복적인 클러스터링 알고리즘인 K-평균 알고리즘 등을 사용하여 구성한 코드북을 이용하여 부호화 하게 된다[2]. 각 입력 벡터는 가장 거리가 가까운 하나의 코드워드로 양자화 되는데 데이터를 전송할 때는 원래의 데이터 대신 이들 코드워드의 인덱스만을 보냄으로써 압축을 이룰 수 있다. 복호화는 전송된 인덱스를 이용하여 이에 해당하는 코드워드의 값으로 단순히 치환함으로써 이루어진다.

벡터양자화에서 핵심적인 사항은 주어진 학습데이터를 이용해서 어떻게 하면 가장 좋은 코드북을 생성할 수 있는 가이다. 코드북을 생성하는 알고리즘 중에서 가장 널리 쓰이는 알고리즘에는 K-평균 알고리즘 또는 Linde, Buzo, Gray가 제안한 K-평균 알고리즘과 본질적으로는 동일한 LBG 알고리즘이 있다. 이 알고리즘은 학습데이터와 이에 대응하는 코드워드 사이의 오차를 반복적으로 감소시켜 나가는 방법을 사용하는데, 전역적으로 최적인 코드북을 구할 수는 없고 국부적으로 최적화된 코드북을 얻을 수밖에 없다는 단점을 지닌다.

따라서 K-평균 알고리즘을 변형시켜 보다 나은 코드북을 얻고자 하는 노력들이 많이 행해졌는데 그 중 수정된

* 서울대학교 전기공학부
접수일자: 1997年 4月 29日

K-평균 알고리즘(Modified K-means Algorithm, 이하 MKM 알고리즘)은 구현이 간단하고 계산량이 많지 않으면서도 좋은 코드북을 얻는 알고리즘이다[2]. MKM 알고리즘은 자기 집단에 속하는 데이터의 중심을 새로운 코드워드로 삼는 것이 아니라 현재 코드워드와 새로 구한 집단의 중심을 연결한 선상에서 새로 구한 중심 너머의 일정한 점을 새로운 코드워드로 선택하는 방식인데, 본 논문에서는 이렇게 구한 새로운 코드워드가 어떠한 조건을 만족할 때 알고리즘이 반복적 감소의 성질을 가지는지 살펴보고, 그 수렴 영역 중 기존의 방식보다 더 좋은 성능을 보이는 코드워드를 선택하는 방법을 제시함으로써 개선된 VQ 학습 알고리즘을 제안하고자 한다.

II. 기존 알고리즘 및 제안된 알고리즘

코드워드 y에 대한 클래스를 $C(y)$라고 정의하고 어떤 학습데이터가 코드북의 모든 코드워드 중에서 코드워드 y와의 거리오차가 가장 작다면 이 학습데이터는 클래스 $C(y)$의 원소가 된다. 클래스 $C(y)$의 중심은 다음과 같이 정의된다.

$$centroid(C(y)) = \frac{1}{|C(y)|}\sum_{x \in C(y)} x$$

여기서 $|C(y)|$는 $C(y)$ 클래스에 속하는 학습데이터의 개수를 나타낸다. 모든 학습데이터와 각각에 해당하는 코드워드들 사이의 평균 거리오차는 다음과 같이 정의된다.

$$D = \frac{1}{|T|}\sum_{x \in T} d(x, \hat{x})$$

여기서 T는 총 학습데이터이고, $\hat{x}$는 x에 가장 가까운 거리에 있는 코드워드이다. 만약 C_m를 m번째 반복시 코드북 그리고 y_i^m을 m번째 반복시 i번째 코드워드라고 하고, 수렴 임계값을 ε라고 하면 MKM 알고리즘은 다음과 같다.

0 단계 : 주어진 초기 코드북을 C_0라고 하고 $m=0$, $D_{-1}=\infty$로 둔다.

1 단계 : 각 학습데이터 x에 대해 C_m 중에서 최단거리에 있는 코드워드를 찾는다.

2 단계 : $y^{m,c} = centroid(C(y^m))$를 구한다.

3 단계 : $y_i^{m+1} = y_i^{m,c} + \text{scale} \cdot (y_i^{m,c} - y_i^m)$ 로부터 C_{m+1}을 얻는다.

4 단계 : 평균오차 D_m를 계산하고, 만약 다음의 판단조건 $(D_{m-1} - D_m)/D_m \le \varepsilon$이 만족되면 멈춘다. 만약 그렇지 않으면 $m=m+1$로 하고 Step 1로 간다.

여기서 1 단계는 부호화 과정이고, 2 단계와 3 단계가 코드북 갱신 과정이다. scale 값은 -1에서 1까지 변화할 수 있는데 왜 이러한 범위 내에서 값이 변화해야 하는지는 뒤에서 반복적 감소성질을 다룰 때 하기로 한다. 위의 3단계를 약간 변화시켜서 식을 다시 써보면 다음과 같다.

$$y_i^{m+1} = y_i^{m,c} + \text{scale} \cdot \triangle y_i \;(\text{여기서}\;\; \triangle y_i = y_i^{m,c} - y_i^m \text{이다}) \tag{1}$$

식을 위와 같이 변화시켜보면 MKM 알고리즘은 K-평균 알고리즘과는 달리 $y_i^{m,c}$를 코드워드로 하지않고 특정 방향 $\triangle y_i$로 더 전진한 어떤 점을 새로운 코드워드로 삼는다는 것을 알 수 있다. 이것을 도시한 것이 그림 1이다. 그림을 보면 새로운 중심점($y_i^{m,c}$)을 코드워드로 삼는 것이 아니라 현재의 코드워드($y_i^{m,c}$)와 새로운 중심점을 연장한 선상에 임의의 점을 새로운 코드워드(y_i^{m+1})로 삼는다는 것이 명확하다. 새로운 코드워드의 정확한 위치는 scale 값이 정해지면 결정된다.

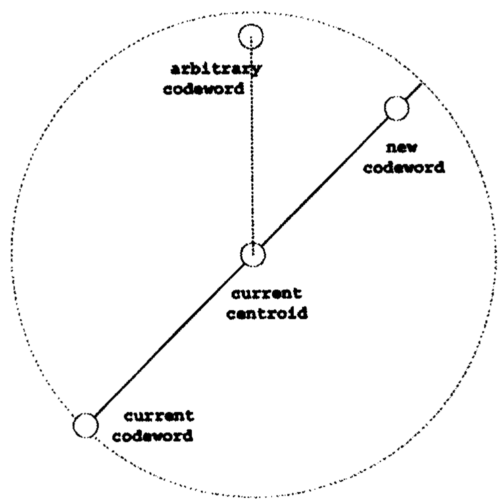

그림 1. 수렴 영역에 대한 도시

식 1에서 생각해야 할 점은 두 가지이다. 첫째는 scale 값을 어떻게 잡는 것이 좋은가 인데 이 문제는 수학적으로 구할 수 없다. 따라서 시료 데이터를 이용해 실험적으로 좋은 scale 값을 취할 수밖에 없는데, 본 논문의 실험에서도 이와 같은 방법을 취하였다. 둘째는 식 1과 같이 K-평균 알고리즘과 다른 점에서 새로운 코드워드를 잡게 될 때 어떠한 영역 내에 그 코드워드를 잡아야 알고리즘이 반복적 감소성질을 만족하는가 이다. 이 수렴범위문제에 대해서는 다음 정리를 살펴보기로 한다.

정리) y_i는 i번째 코드워드이고, 전체 코드북의 크기는 이라고 하며, m번째 코드북 $Y^m = \{y_i^m, 1 \le i \le K\}$ 에 의한 Mean-Squared-Error(이하 MSE)를 $D(m)$이라 하고, m번째 반복과정에서 i번째 집단의 중심을 $y_i^{m,c}$라고 하자. 이때 $m+1$번째 코드북 $Y^{m+1} = \{y_i^{m+1}, 1 \le i \le K\}$의 코드워드가 각각 다음의 성질을 만족하면

$$|y_i^{m}-y_i^{m+1}|^2 \ge |y_i^{m+1}-y_i^{m,c}|^2, \quad (2)$$

동일한 분할에 대하여 $m+1$번째 코드북에 의한 MSE $D(m+1)$은 $D(m+1) \ge D(m)$을 만족한다[3].

증명) 전체 MSE를 D라 하고, w_{ij}는 membership 함수로 i 번째 데이터가 j번째 코드워드에 속할 때만 1이고 나머지는 0인 함수라 하면 D는 다음과 같이 쓸 수 있다.

$$\begin{aligned} D &= \sum_i \sum_{j=1}^{K} w_{ij} |x_i-y_j|^2 \\ &= \sum_{j=1}^{N_1} |x_i-y_1|^2 + \sum_{i=N_1+1}^{N_2} |x_i-y_2|^2 + \cdots + \sum_{i=N_{K-1}+1}^{N_K} |x_i-y_K|^2 \\ &= D_1 + D_2 + \cdots + D_K \end{aligned}$$

여기서 x_i는 i번째 입력 데이터이며, N_i는 i번째 군집까지의 데이터 개수이다. 이 중 $D_1(m)-D_1(m+1)$을 적어보면

$$\begin{aligned} D_1(m)-D_1(m+1) &= \sum_{i=1}^{N_1} |x_i-y_1^{(m)}|^2 - \sum_{i=1}^{N_1} |x_i-y_1^{(m+1)}|^2 \\ &= \sum_{i=1}^{N_1} (|y_1^{(m)}|^2 - |y_1^{(m+1)}|^2 - 2x_i \cdot y_1^{(m)} + 2x_i \cdot y_1^{(m+1)}) \\ &= N_1(|y_1^{(m)}|^2 - |y_1^{(m+1)}|^2) - 2((y_1^{(m)}-y_1^{(m+1)}) \cdot \sum_{i=1}^{N_1} x_i) \end{aligned}$$

이 되고 첫 번째 집단의 중심 y_i^c는 $\sum_{i=1}^{N_1} x_i$ 이므로

$$D_1(m)-D_1(m+1) = N_1(|y_1^{m}|^2 - |y_1^{m+1}|^2) - 2N_1((y_1^{(m)}-y_1^{(m+1)}) \cdot y_1^{m,c})$$

이 된다. 이것을 다시 완전 제곱으로 나타내면

$$D_1(m)-D_1(m+1) = N_1(|y_1^{m}-y_1^{m,c}|^2 - |y_1^{m+1}-y_1^{m,c}|^2)$$

이 되는데 식 2를 위의 식에 적용하면 0보다 크게 되어 $D_1(m+1) \le D_1(m)$가 성립된다. 이와 같은 관계를 모든 집단에 대하여 확장하면 $D(m+1) \le D(m)$이 성립하는 건 자명하다.

위에서 본 것처럼 새롭게 구한 코드벡터 y_i^{m+1}가 식 2를 만족하면 코드북 갱신과정은 반복적 감소성질을 만족하게 되는데 이것을 수정된 K-means 방법에 적용하면 scale 값이 -1에서 1까지 변화할 때라는 걸 알 수 있다. 이 영역을 도시한 것이 그림 1의 점선으로 된 원이다.

이상을 살펴볼 때 Δy_i는 수렴원의 중심에서 아무 방향으로나 수렴원의 크기보다 작게 나아가기만 하면 반복적 감소성질이 만족되므로 MKM에서 제시한 직선 외에 다른 곳에 새로운 코드워드를 잡을 수 있다. 한 예를 그림 1에 점선으로 임의의 코드워드 형태로 나타내었다. 본 논문에서는 MKM 알고리즘의 3단계를 다음과 같이 수정하였다.

3 단계 : $y_i^{m+1} = y_i^{m,c} + \text{scale} \cdot \frac{|y_i^{m,c}-y_i^{m}|}{|y_i^{m,c}-y_i^{m-1}|}(y_i^{m}-y_i^{m-1})$에서 C_{m+1}을 얻는다.

이 방법에서는 $\Delta y_i = y_i^{m}-y_i^{m-1}$로 잡았는데, scale 바로 뒤에 붙은 식은 수렴영역 안으로 Δy_i가 들어오도록 하는 정규화식이다. MKM은 현재 반복 때의 정보만을 Δy_i에 이용하므로 이전 반복 때의 정보를 이용하면 좀 더 나은 코드북을 얻을 수 있지 않을까 하는 착상이 위의 갱신식에 깔린 생각이다. 이러한 착안은 Δy_i를 경사도(gradient)로 생각하면 일반 adaptive signal processing에서 쉽게 찾아볼 수 있다[4]. 새롭게 수정된 3단계 갱신 식의 유효성은 실험을 통해서 확인해 볼 수 있는데, 제안된 방법을 사용할 때 맨 처음 갱신에서는 y_i^{-1}을 얻을 수 없으므로 MKM과 동일한 갱신식을 사용하여 알고리즘이 구성되어야 한다는 점을 잊지 말아야 한다.

III. 실험 및 토론

제안된 알고리즘의 성능을 평가하기 위해서 4개의 USC 화상데이터를 이용하여 실험하였다. 화상데이터는 각각 512 x 512 흑백 정지 화상이며 각 화상의 화소는 256개의 흑백 단계로 이루어져 있다. 본 실험에서 벡터의 차원은 4x4=16을 사용했으며, 제안된 알고리즘은 학습 속도를 빠르게 하기 위해서 고속 부호화 방법인 ENNS(Equal average Nearest Neighbor Search)를 이용하였는데, 이 방법은 전체 탐색 알고리즘과 동일하며 단지 속도만 빠르게 하는 부호화 방법이다[5][6]. 각 실험에서 사용된 초기 코드북은 각 화상으로부터 binary split 알고리즘을 이용하여 구성되었으며 실험에 사용한 수렴 임계값은 0.0005로 정했다.

먼저 적절한 scale 값을 정하기 위해 scale 값을 MKM 알고리즘과 제안된 방법에 대해서 변화시켜가며 얻은 코드북의 전체 에러값을 조사하였다. 이때 사용한 화상은 Lena이며 코드북의 크기는 256인데, MKM 알고리즘에서 결과가 표 1에 나와있다. 표 1을 보면 MKM 알고리즘의 경우에는 scale 값이 커질수록 좀더 좋은 코드북을 얻을 수 있으나 0.97이후로는 코드북이 다시 나빠지는 것을 확인할 수 있다. 또한 반복 횟수의 경우에는 scale이 0.80을 중심으로 좌우로 갈수록 커지는 것을 확인할 수 있다. 속도가 중시되는 경우는 scale 값이 0.80인 경우를 선택하는 것이 좋겠지만 학습의 가장 중요한 목표가 좋은 코드북의 생성에 있으므로 본 논문에서는 scale 값을 0.95로 선택하여 사용하였다.

표 1. MKM 알고리즘에서 scale 변화에 따른 MSE 및 반복횟수

scale	0.7	0.75	0.80	0.85	0.90	0.95	0.97	0.99	0.995
MSE	41.96	41.94	41.95	41.94	41.83	41.72	41.76	41.76	41.87
반복횟수	27	25	22	23	35	48	51	92	109

제안된 알고리즘의 경우는 표2에 결과를 표시하였다. 표 2도 표 1과 동일한 경향을 보이는데 가장 좋은 코드북을 생성해내는 scale 값이 0.97이라는 점에서만 차이가 있다. 따라서 본 실험에서 제안된 방법을 적용할 때는 가장 좋은 코드북을 생성한 scale 값 0.97을 사용하였다.

표2. 제안된 알고리즘에서 scale 변화에 따른 MSE 및 반복횟수

scale	0.7	0.75	0.80	0.85	0.90	0.95	0.97	0.99	0.995
MSE	42.16	41.82	41.84	41.81	41.72	41.69	41.46	41.47	41.50
반복횟수	31	29	27	35	34	43	55	81	85

표 3은 사용된 네 가지 화상에 대해서 코드북의 크기를 128부터 512까지 변화시키면서 구한 MSE이다. 이 표에서는 Baboon 화상 코드북 크기 256일 경우를 제외하고는 다른 모든 경우에 새로 제안된 방법이 기존의 방법에 비해 MSE의 관점에서 더 좋은 코드북을 생성해 낸다는 것을 확인할 수 있다. 따라서 제안된 방법이 효과적이라는 것은 분명하다.

표3. 실험 화상에 대한 MSE

크 기	방 법	화 상			
		Lena	Kan	Boats	Baboon
128	K-평균	54.60	87.52	112.4	289.5
	MKM	53.94	86.73	111.4	288.0
	제안방법	53.64	86.32	111.1	287.1
256	K-means	42.66	71.08	92.25	249.9
	MKM	41.72	69.72	90.83	245.9
	제안방법	41.46	69.37	90.40	246.3
512	K-평균	33.13	57.21	75.23	211.8
	MKM	31.92	55.36	72.98	207.6
	제안방법	31.53	54.65	72.12	206.3

IV. 결 론

본 논문에서는 기존에 제시된 MKM 알고리즘에 대해서 분석하였다. MKM 알고리즘에서 사용된 코드북 갱신과정에서 새로운 코드워드가 어떠한 조건을 만족할 때 알고리즘이 반복적 감소의 성질을 가지는지 살펴보았으며, 그 수렴 영역 중 기존의 방식보다 *MSE* 관점에서 더 나은 성능을 보이는 코드워드를 선택하는 일례를 제시하였고 이를 실험으로 확인하였다. 하지만 본 논문에서 제시한 방식보다 더 좋은 코드북 갱신식이 나올 수 있으므로 가장 좋은 갱신식이 어떠한 것일까에 대해서는 좀 더 연구가 필요할 것이다.

참 고 문 헌

1. R. M. Gray, "Vector Quantization," IEEE ASSP Magazine, vol. 1, pp. 4-9, Apr. 1984.
2. D. Lee, S. Baek, and K.-M. Sung, "Modified K-means Algorithm for Vector Quantizer Design," IEEE Signal Processing Letters, vol. 4, No. 1, pp. 2-4, Jan., 1997.
3. 이대룡, "개선된 부호책 갱신 방법과 중첩 집단화 방법을 이용한 벡터양자화기의 부호책 구성 알고리듬," 서울대학교 박사학위논문, 1997.
4. S. Haykin, Adaptive Filter Theory, Prentice Hall, 1996.
5. S. W. Ra and J. K. Kim, "A fast mean-distance-ordered partial codebook search algorithm for image vector quantization," IEEE Trans. Circuits Syst. II, vol 40, no. 9, pp. 576-579, Sept. 1993.
6. S. Baek, B. Jeon, and K.-M. Sung, "A Fast Encoding Algorithm for Vector Quantization," IEEE Signal Processing Letters, vol. 4, No. 12, pp. 325-327, Dec., 1997.

▲백 성 준

한국음향학회지 16권 제2호 참조

현재 : 서울대학교 대학원 전기공학부 박사과정

▲최 용 진 1965년 11월 14일생

1988년 2월 : 서울대학교 전자공학과 졸업(공학사)

1990년 2월 : 서울대학교 전자공학과 졸업(공학석사)

1990년 3월~현재 : 서울대학교 전기공학부 박사과정

▲이 주 헌 1964년 6월생

1990년 2월 : 서울대학교 전자공학과 졸업(공학석사)

1995년 2월 : 서울대학교 전자공학과 졸업(공학박사)

1997년 3월~현재 : 동아방송대학 정보통신과 조교수로 재직중

※주관심분야 : 음성신호처리, 멀티미디어 및 통신시스템

▲성 굉 모

현재 : 서울대학교 전기공학부 교수