

논문-99-4-1-05

반복적 병합을 이용한 윌래곡 기반 영상 분할

전성철*, 임채환*, 김남철*

Rate-Distortion Based Image Segmentation Using Recursive Merging

Sung Chul Jun*, Chae Whan Lim*, and Nam Chul Kim*

요약

본 논문에서는 영역 인접성 그래프를 사용한 반복적 병합법으로 윌래곡 기반 영상 분할을 수행하는 알고리즘을 제안하였다. 제안한 기법에서는, 인접한 두 영역간의 이질성을 윌래곡적 관점에서 고려한 Lagrange 비용함수로 표현하였다. 이때 각 병합 단계에 맞는 Lagrange 상수를 추정하고, 윌래곡 기반 비용이 최소가 되는 인접 영역 쌍을 탐색하여 이들을 새로운 영역으로 병합하였다. 이러한 병합과정은 어떤 정지조건에 이를 때까지 반복한다. 제안된 분할법은 영역기반 부호화 혹은 분할기반 부호화 기법에 적합한 분할법이다. 실험 결과, 256x256 Lena에 대해서 영역기반 부호화를 수행하였을 때, 제안한 기법은 평균 밝기 차이, 윌래곡의 변화량 및 JPEG에 비하여 각각 2.5~3.5 dB, 0.8~1.0 dB, 0.3~0.6 dB 정도의 PSNR 성능 개선이 있는 것으로 나타났다.

Abstract

In this paper, a rate-distortion based image segmentation algorithm is presented using a recursive merging with region adjacency graph (RAG). In the method, the dissimilarity between a pair of adjacent regions is represented as a Lagrangian cost function considered in rate-distortion sense. Lagrangian multiplier is estimated in each merging step, a pair of adjacent regions whose cost is minimal is searched and then the pair of regions are merged into a new region. The merging step is recursively performed until some termination criterion is reached. The proposed method thus is suitable for region-based coding or segmented-based coding. Experiment results for 256x256 Lena show that segmented-based coding using the proposed method yields PSNR improvement of about 2.5~3.5 dB, 0.8~1.0 dB, 0.3~0.6 dB over mean-difference-based method, distortion-based method, and JPEG, respectively.

I. 서론

영상 분할(segmentation)^{[1][2][3]}은 영상 분석, 물체 인식 등의 컴퓨터 비전을 위한 기본 과정으로 사용되어 왔으며, 최근에는 동영상 표준기법인 MPEG-4^[4]에서 어떤 장면을

그 장면을 구성하고 있는 기본 객체를 별로 표현하기 위한 방법으로 사용되고 있어, 그 중요성이 강조되고 있다. 영상 분할은 입력영상을 동질성(homogeneity)과 연결성(connectivity)을 만족하는 영역들로 나누는 것으로 정의할 수 있다. 여기서 동질성은 영역내의 모든 화소들이 공통의 성질을 갖는 것이고, 연결성은 영역내의 임의의 두 화소사이에 연결된 경로(path)가 존재하는 것을 나타낸다.

지금까지 연구된 영상 분할법들은 군집화(clustering), 에지기반법(edge-based method), 영역기반법(regionbased method), 그리고 이들 방식을 결합시킨 하이브리드법(hybrid method) 등으로 크게 분류할 수 있다. 이들 중에

* 경북대학교 전자·전기공학부

School of Electronic & Electrical Eng., Kyungpook Nat'l University.

* 본 연구는 한국과학재단 '97특정연구 지원(97-0100-02-01-3)에 의하여 수행하였음.

서 영역기반법이 세밀한 분할이 가능하기 때문에 분할 결과가 다른 방식에 비해서 우수한 것으로 알려져 있으며, 영상부호화 분야에서, 분할 영상을 영역의 내부 질감(texture)과 외부 윤곽(contour)으로 나누어 부호화 하는 영역기반 부호화 기법 (또는 분할기반 부호화 기법)에 적용하고자 할 때, 주로 이러한 영역기반 분할법을 이용하고 있다^{[5][6]}.

이러한 영역기반법에는 분할법(split-and-merge)^[3], CLRG(centroid linkage region growing)^[7], RAG(region adjacency graph)를 이용한 분할법^{[6][8][9][10]} 등이 있다. 분할법은 우선 쿼드트리(quad tree) 형태의 기본적인 구조 내에서 영역을 분할하거나 병합한 후, 나중에 쿼드트리 형태를 깨면서 인접한 영역들간의 병합(merging)을 수행하는 것으로 비교적 구조가 간단하고 수행 시간이 매우 짧지만, 많은 경우 분할 영상에는 정방형 형태의 영역 윤곽들이 그대로 남아 있어 눈에 거슬리게 된다. CLRG는 주사(scanning) 순서에 따라 영역을 병합하기 때문에 구조가 간단한 반면 영상내의 전체적 정보를 제대로 이용하지 못하며 분할 결과가 주사 순서에 따라 달라진다. 또한, RAG를 이용한 분할법은 그래프 이론을 영상 분할에 적용시킨 것들이다. 이들은 초기 분할된 영상으로부터 정점(vertex), 가지(edge) 그리고 동질성의 조건이 되는 비용(cost)을 부여한 가중치 그래프(weighted graph)를 구성한 후 최소 가중치 가지를 탐색하여 원하는 수준에 이를 때까지 반복적으로 병합하는 방식이다.

이러한 영역기반 분할법들에서는, 특히 인접한 영역들 사이의 동질성 및 이질성(dissimilarity)을 어떻게 정의하느냐에 따라 분할 영상의 특성이 달라진다. 따라서, 분할이 적용되는 응용분야에 맞게 동질성 및 이질성에 대한 기준 혹은 척도가 달라야 한다. 어떤 장면에서 독립적인 객체를 분할하기 위한 것인지 아니면 부호화에 적용하기 위한 분할인지를 고려해서 다양한 동질성 및 이질성의 척도, 즉 비용을 정의할 수 있다.

기존의 영역기반 분할법들에서는 인접한 두 영역간의 이질성을 측정하는 비용으로 두 영역간의 평균밝기 차이(difference of mean intensity)^{[9][10]}, 두 영역간의 왜곡의 변화량(difference of distortion)^[5], 그리고 윌쇼프 기울기(rate-distortion slope)^[11] 등을 사용한 바 있다. 그러나, 이 비용들은 윌쇼프 이론에 근거하여 분할 영상의 부호화 효율을 극대화하도록 고려된 것은 아니다.

본 논문에서는 분할 영상의 부호화 효율을 극대화하기 위하여, 인접 영역간의 이질성의 비용으로 윌쇼프 기반

(rate-distortion based) 비용을 정의하고, 이 비용에 근거하여 RAG를 이용한 반복적 병합법(recursive merging)으로 영상 분할을 수행하였다. 제안한 윌쇼프 기반 비용은 최적의 조건을 만족하는 두 영역을 병합하기 위하여 Lagrange 상수를 이용해서 각 단계에서의 윌 및 왜곡을 동시에 고려한 것이다. 그리고, 이러한 윌쇼프 기반 분할 알고리즘의 부호화 성능을 조사하기 위하여, 분할 영상들을 영역의 내부 질감은 평균밝기로 표현하여 고정길이로 부호화하고, 영역들의 윤곽은 윤곽화소들을 따라가면서 부호화 하는 여섯 방향 윤곽부호화^[12]로 부호화 된다. 실험에서는, 제안된 윌쇼프 기반 비용과 기존의 비용에 따른 분할 결과들을 영역기반 부호화 기법을 사용하여, 부호화 비트율에 대한 PSNR로써 부호화 성능을 조사하고, 정지영상 부호화의 표준 기법인 JPEG과도 그 성능을 비교한다.

II. 반복적 병합에 의한 영상 분할

1. 영상 분할의 정의

영상 분할은 입력영상 I 를 동질성 및 연결성을 만족하는 영역들로 나누는 것으로, 분할된 결과는 파티션(partition) $P = \{R_1, R_2, \dots, R_N\}$ 로 나타낼 수 있다. 여기서 $R_i \in P$ 인 R_i 들은 동질성 및 연결성을 만족한다. 분할된 영상 P 는 I 의 파티션이므로, $R_i \cap R_j = \emptyset$ (단, $i \neq j$)와 $\bigcup_k R_k = I$ 를 만족한다. 본 논문에서는 여러 가지 영역기반 분할법 중에서 RAG를 이용한 반복적 병합법을 선택하여 영상 분할을 수행하기로 한다.

2. RAG를 이용한 반복적 병합

초기 분할된 영상의 각 영역을 그래프의 정점으로 대응시키고, 인접한 영역 쌍 및 그들의 이질성 정도를 그래프의 가지 및 그 가지의 비용으로 대응시키면 RAG를 얻을 수 있다. 예를 들어, 초기에 영상이 화소 단위로 분할되었다면, 그림 1(a)의 4x4 시험 영상(test image)에 대하여 화소들 사이의 4방향 연결성을 고려하여 대응시킨 RAG는 그림 1(b)와 같이 16개의 정점과 24개의 가지들로 구성된다. 만약, $N \times N$ 크기의 영상이라면 초기의 정점은 $N \times N$ 개, 가지는 $2N(N-1)$ 개를 가진다. 이렇게 구성된 그래프를 이용하여 모든 가지에 대하여 비용을 부여하고, 매 단계마다 최소 비용을 갖는 가지를 선택하여 연결된

두 영역씩 병합해간다.

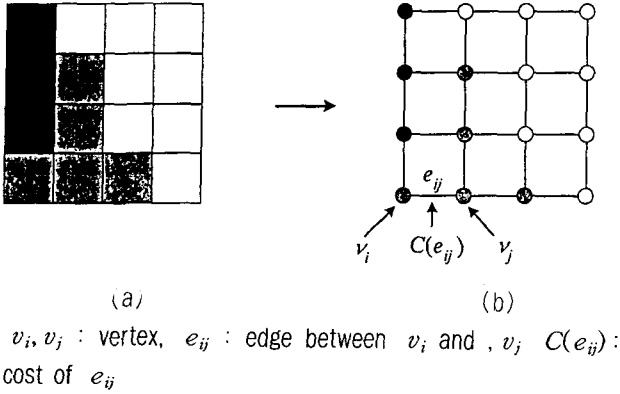


그림 1. 영상에서 그래프로의 사상 (a) 시험영상, (b) RAG
Fig. 1. Mapping image onto Graph: (a) test image and (b) RAG.

K 개의 영역들로 나누어진 $P_t = \{R_1, R_2, \dots, R_K\}$ 는 병합 단계 t 에서 얻어진 최적의 파티션이라고 하자. 여기서 t 의 범위는 $0 \leq t \leq N^2 - 1$ 이다. 그러면, P_t 에 대응하는 RAG는 $G(P_t) = (V_t, E_t)$ 로 구성된다. 여기서 V_t, E_t 는 각각 정점들과 가지들의 집합을 나타낸다. 즉, $R_i \in P_t$ 이면 $v_i \in V_t$ 이고, R_i 와 R_j 가 인접한다면 $\{v_i, v_j\} \in E_t$ 가 된다. 단, $0 \leq i, j \leq K, i \neq j$ 이다. 구성된 그래프에서 $e \in E_t$ 인 모든 가지에 대하여 인접한 영역간 이질성의 비용 $c(e)$ 를 부여한다. 이 비용은 인접 영역간의 이질성 정도 및 공유하는 윤곽의 특성 등을 고려하여 부여할 수 있다. 다음은, $G(P_t)$ 의 가지 중

$\{v_k, v_l\} = \arg \min_{e \in E_t} \{c(e)\}$ 를 만족하는 최소 비용 가지를 찾는다. 그러면, 이 가지의 두 정점 v_k, v_l 에 대응하는 영역 R_k 와 R_l 을 병합하여 새로운 영역 $R_k \cup R_l \in P_{t+1}$ 을 형성하게 된다. 한 단계 병합한 결과, $K-1$ 개의 영역으로 이루어진 새로운 파티션 $P_{t+1} = \{R_1, R_2, \dots, R_k \cup R_l, \dots, R_K\}$ 를 얻게 되며, P_{t+1} 에 대응하는 $G(P_{t+1}) = (V_{t+1}, E_{t+1})$ 로 구성된 RAG를 구할 수 있다. 여기서 V_{t+1} 는 $V_{t+1} = V_t \cup \{v_{kl}\} - \{v_k, v_l\}$ 이고, E_{t+1} 는 어떤 n 에 대하여 $\{v_k, v_n\} \in E_t$ 와 $\{v_l, v_n\} \in E_t$ 일 때 $E_{t+1} = E_t - \{\{v_k, v_l\}\} - \bigcup_n \{\{v_l, v_n\}\}$ 으로 나타낼 수 있다. 앞에서 v_{kl} 는 병합된 새로운 영역의 정점을 나

타낸다.

이렇게 최소 비용 가지에 연결된 두 영역을 병합한 후 병합된 영역과 연관된 가지들의 비용을 다시 갱신한다. 최소 비용의 영역 쌍이 병합되는 과정을 그림 2에 나타내었다. 이러한 과정을 정지 조건에 도달할 때까지 반복한다. 일반적으로 정지조건은 정해진 영역 수 혹은 허용된 왜곡 수준이 될 수 있다. 만약 정지조건이 영역 수일 경우, K 개의 초기 영역이 주어졌을 경우 원하는 M 개의 분할 영역을 얻기 위해서는 단계적 병합을 $K-M$ 번 (단, $M \leq K$) 반복하면 된다. 이러한 RAG를 이용한 분할 과정을 그림 3에서 블록도로 나타내었다.

3. 기존의 가지 비용

영역기반 분할법의 가지 비용은 인접한 두 영역간의 이질성을 나타내는 척도이며, 분할 결과를 결정하는 가장 중요한 요인으로, 기존의 분할법에서는 평균밝기 차이, 왜곡의 변화량 또는 울왜곡 기울기 등을 이용하였다. 즉, 인접한 두 영역 R_i 과 R_j 간의 이질성에 관한 비용 $c(\{v_i, v_j\})$ 은 다음과 같이 정의되어 사용되었다.

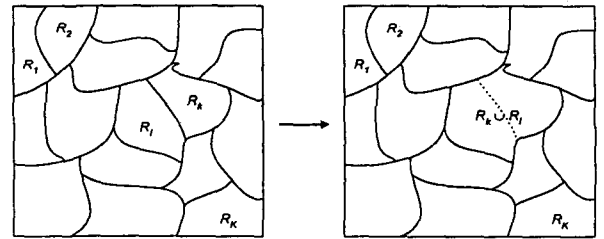


그림 2. 한 단계 병합 과정
Fig. 2. Example of one-step recursive merging.

평균밝기 차이^{[9][10]} :

$$c(\{v_i, v_j\}) = \Delta M = |\mu(R_i) - \mu(R_j)| \quad (1)$$

여기서 $\mu(R_i)$ 는 영역 R_i 의 평균밝기이다.

왜곡의 변화량^[5] :

$$c(\{v_i, v_j\}) = \Delta D = D(R_i \cup R_j) - (D(R_i) + D(R_j)) \quad (2)$$

여기서 $D(R_i) = \sum_{(x,y) \in R_i} [I(x,y) - \hat{I}(x,y)]^2$ 로서 누적자승오차(sum of squared error : SSE)의 변화량을 나타낸다. $I(x,y)$, $\hat{I}(x,y)$ 는 각각 원 영상과 근사화된 영상의 밝기를 나타내고, x,y 는 화소위치를 나타낸다.

울왜곡 기율기^[11] :

$$c(\{v_i, v_j\}) = |\Delta D| / |\Delta R| \quad (3)$$

여기서 왜곡의 변화량 ΔD 은 (2)식과 같고, 울의 변화량 ΔR 은 두 영역이 한 영역으로 병합되면서 발생하는 윤곽 및 질감에 소요되는 비트(bits)의 변화량이다. 그림 1에서 보면, 병합되면서 발생하는 울의 변화량은 두 영역 R_i 와 R_j 가 공유하는 윤곽화소들이 사라지면서 감소되는 비트와 두 영역을 하나의 영역으로 표현될 때 감소되는 비트의 합으로 주어진다.

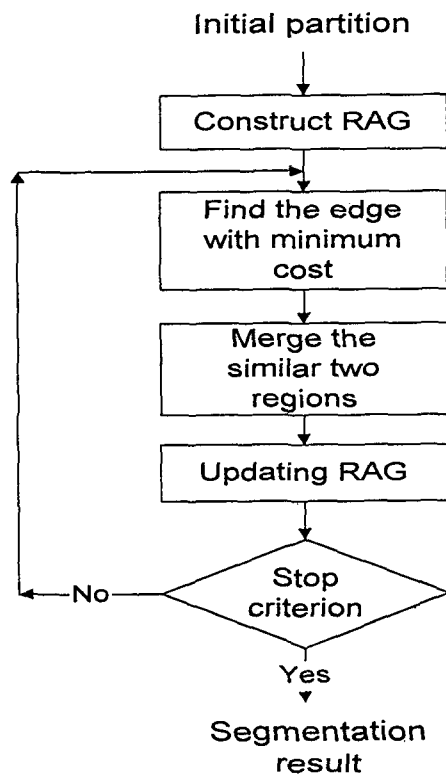


그림 3. RAG를 이용한 반복적 병합법의 블록도
Fig. 3. Block diagram of recursive merging method using RAG.

평균밝기 차이에 의한 분할에서는 밝기 값을 기준으로 두 영역간의 이질성의 정도를 판단하여 밝기 변화가 최소인 가지를 병합하는 것이다. 비용의 계산량은 적으나, 영상에 존재하는 잡음의 영향에 매우 민감하기 때문에 분할 결과에서 소 영역이 무수히 발생하여 이를 억제하기 위한 전처리 과정 혹은 후처리 과정이 요구된다. 그러나, 비용 계산이 매우 간단하기 때문에 지금까지도 여러 분할 알고리즘에서 많이 사용되고 있다.

왜곡의 변화량에 의한 분할에서는 SSE의 변화량을 가지의 비용으로 사용하기 때문에, 왜곡의 증가량이 가장 적게 발생시키는 가지를 우선 병합한다. 따라서, 주어진 영역 수에 대하여 평균자승오차(mean squared error : MSE)가 최소인 분할 영상을 얻을 수 있으므로, 최대의 PSNR을 가진다. 그러나, 분할 영상의 부호화에 소요되는 비트를 고려하지 않았으므로 주어진 비트율에 대하여 최적의 PSNR을 가지는 것은 아니다.

마지막으로, 울왜곡 기율기를 이용한 분할에서는 각 단계에서 울과 왜곡의 변화량을 조사하여 그 기율기가 최소인 가지 즉, 단위 비트당 최소 왜곡을 발생시키는 가지에 연결된 두 영역을 병합시키는 것이다. 이것은 울왜곡적 관점에서 깊게 해석되어 제시된 것은 아니지만, 울과 왜곡을 동시에 고려한 비용을 사용함으로써 기존의 다른 비용을 사용한 방법에 비해 성능 개선이 있을 것을 기대할 수 있다.

III. 울왜곡 기반 영상 분할

1. 울왜곡 기반 비용

영역기반 부호화에 적합한 분할이 된다는 것은, 분할 영상을 부호화하였을 때 주어진 비트 R 에 대하여 왜곡 D 가 최소가 되어야 함을 의미한다. 이러한 문제는 Lagrange 상수 λ 를 이용하여 Lagrangian 비용함수 $J = D + \lambda R$ 을 최소화하는 문제로 생각할 수 있다^[13]. 본 논문에서는 병합의 매 단계마다 주어진 파티션에 대해서 부호화 관점에서의 Lagrange 비용함수를 정의하고 이에 의한 단계별 최적의 분할을 수행해 감으로써 국부 최적의 분할결과를 얻고자 한다.

우선, 반복적 병합의 t 단계에서 파티션 P_t 가 최적화 되었다고 가정하고, 이때의 P_t 에 대한 Lagrangian 비용함수 $J(P_t)$ 는 다음과 같다고 하자.

$$J(P_t) = D(P_t) + \lambda_t R(P_t) \quad (4)$$

여기서 λ_t 는 t 단계에서의 Lagrange 상수이다. 그러면, 가지 $e \in E_t$ 를 병합했을 때 형성되는 파티션 P_{t+1} 에 대한 비용함수는 다음과 같이 주어진다.

$$J(P_{t+1}) = J(P_t) + \Delta J(e) \quad (5)$$

여기서 $\Delta J(e)$ 는 가지 e 의 율왜곡 기반 비용 $c(e)$ 을 나타낸다. (5)식으로부터 최적의 파티션 P_{t+1} 을 얻기 위해서는 $\Delta J(e)$ 가 최소화되어야 함을 알 수 있다.

율왜곡 기반 비용 $\Delta J(e)$ 은 (4) 및 (5)식으로부터 t 단계에서 $\{v_i, v_j\} \in E_t$ 를 만족하는 각 가지에 대하여 다음과 같이 정리된다.

$$\begin{aligned} \Delta J(e) &= c(v_i, v_j) \\ &= \{D(P_{t+1}) + \lambda_{t+1}R(P_{t+1})\} - \{D(P_t) + \lambda_t R(P_t)\} \\ &= \Delta D(e) + \lambda_{t+1} \Delta R(e) + (\lambda_{t+1} - \lambda_t)R(P_t) \\ &= \Delta D(e) + \lambda_{t+1} \Delta R(e) + C \end{aligned} \quad (6)$$

여기서 $\Delta D(e)$, $\Delta R(e)$ 은 가지 e 에 연결된 두 영역 사이의 왜곡의 변화량 및 율의 변화량을 나타내고, C 는 $C = (\lambda_{t+1} - \lambda_t)R(P_t)$ 로서 모든 가지에 대하여 동일한 값 가지므로 상수로 나타낸 것이다. 그리고, λ_{t+1} 는 $t+1$ 단계에서의 Lagrange 상수이다.

이상으로, 반복적 병합의 현재 단계에서의 파티션이 P_t 가 최적이라고 가정했을 때, 다음 단계에서의 최적의 파티션 P_{t+1} 을 형성하기 위해서는, t 단계에 존재하는 각 가지에 대하여 (6)식의 율왜곡 기반 비용을 조사하여 그 값이 최소가 되는 가지를 병합하면 된다.

2. Lagrange 상수 추정

위의 (6)식으로부터, t 단계에서의 파티션 P_t 을 고려하여 최적의 조건을 만족하는 가지를 선택하려면, Lagrange 상수 λ_{t+1} 을 추정해서 율왜곡 기반 비용을 계산해야 한다. 일반적으로, 율왜곡의 최적화 문제에서 사용되는 Lagrange 상수는 그림 4와 같이 율왜곡 곡선상의 기울기로 주어진다^[13]. 따라서, t 단계에서의 Lagrange 상수 λ_t

를 알고 있다면, 파티션 P_t 에 대한 비용함수 $J(P_t)$ 는 그림 4와 같이 λ_t 의 기울기를 가지면서 율왜곡 곡선과 접하는 직선이 D 축과 만나는 점의 값이다.

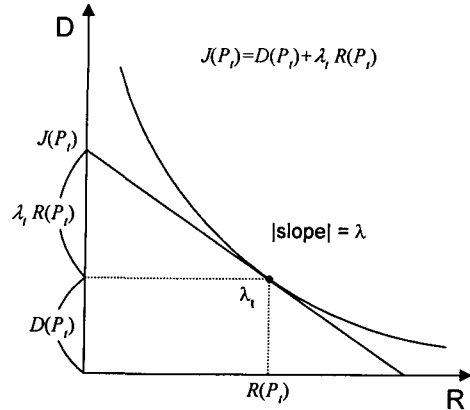


그림 4. 율왜곡 관점에서의 Lagrange 상수 및 비용함수의 의미
Fig. 4. Meaning of Lagrangian multiplier and function in rate-distortion sense.

(6)식의 율왜곡 곡선상의 기울기 λ_{t+1} 은 반복적 병합법의 특성을 고려하여 다음과 같이 추정될 수 있다. 즉, t 단계의 파티션 P_t 에 남아있는 $e \in E_t$ 인 가지들에 대한 율왜곡 기울기가 최소가 되는 값을 사용하여 추정할 수 있다.

$$\lambda_{t+1} = \min_{e \in E_t} \left[-\frac{\Delta D(e)}{\Delta R(e)} \right] \quad (7)$$

이 식은, (3)식과 같이 이 값을 인접한 두 영역간의 이질성의 비용으로 사용한 것^[11]과는 달리, 율왜곡 곡선상의 기울기를 추정하는 데 사용한 것이다. 이와 같이 (7)식으로 Lagrange 상수를 추정할 경우 율왜곡 기반 비용을 해석하면 다음과 같은 중요한 사실을 알 수 있다. 그림 5에서는 현재 파티션 P_t 에서 각 가지를 제거함으로써 P_t 에 대한 율왜곡 평면상의 점에서 P_{t+1} 에 대한 각 점들로 이동되는 것을 몇 개의 가지들에 대하여 예를 들어 나타내었으며, (7)식과 같이 최소의 기울기를 λ_{t+1} 로 추정하였을 때 이들 각 가지에 대한 Lagrangian 비용함수 $J(P_{t+1})$ 를 나타내었다. 이 그림에서 보면, 최소 비용함수 $\min J(P_{t+1})$ 는 λ_{t+1} 을 추정하는 데 사용했던 가지를 제거할 때임을 알 수 있다. 따라서, (7)식으로 Lagrange 상

수를 추정할 경우 율왜곡 기반 비용의 최소값을 갖는 가지는 율왜곡 기울기 기반 비용의 최소값을 갖는 가치와 같음을 알 수 있다.

$$\arg \min_{e \in E_t} \{ \Delta J(e) \} = \arg \min_{e \in E_t} \left\{ - \frac{\Delta D(e)}{\Delta R(e)} \right\} \quad (8)$$

이러한 사실로부터, 기존의 율왜곡 기울기 기반법은 율왜곡 기반법에서 λ_{t+1} 을 (7)식으로 추정하는 하나의 특수한 경우임을 알 수 있다.

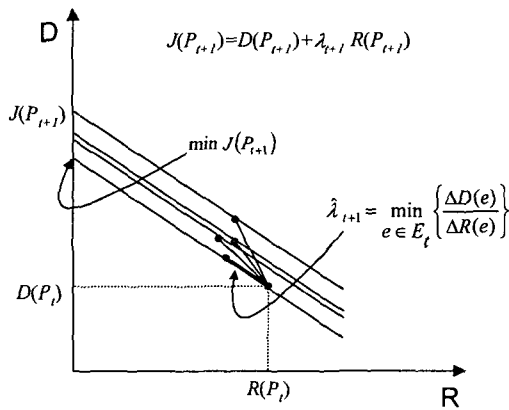


그림 5. 현재 단계에서의 최소 율과 왜곡의 변화량 비를 이용한 Lagrange 상수 추정법
Fig. 5. Lagrangian multiplier be estimated as minim rate-distortion slope in current step.

Lagrange 상수는 병합이 될수록 왜곡에 대하여 단조 증가하는 특성을 가지고 있다.

그런데, (7)식과 같이 그 값을 추정하면, 실제 율왜곡의 기울기에 비해 보다 작은 값으로 추정될 수도 있다. 따라서, 본 논문에서는 단조 증가 특성을 이용하여 이러한 과소 추정을 방지할 수 있도록 다음과 같이 λ_{t+1} 를 추정하기로 한다.

$$\lambda_{t+1} = \max \left[\min_{e \in E_t} \left[- \frac{\Delta D(e)}{\Delta R(e)} \right], \lambda_o, \lambda_t \right] \quad (9)$$

여기서 λ_t 는 이전 단계에서 추정한 Lagrange 상수이고, λ_o 는 병합의 초기 단계에서 왜곡의 변화량이 없는 가치들(이 경우 $\lambda_{t+1} = 0$)에 대해서라도 율의 변화가 고려되도록 하기 위한 상수이다. 또한, 병합 초기 단계에서

형성되는 율왜곡선이 보다 평탄하게 형성되도록 하는 효과도 있다.

3. 율 및 왜곡의 변화량 계산

임의의 가치 e 에 연결된 두 영역을 병합하면서 발생하는 왜곡의 변화량 $\Delta D(e)$ 는 질감 근사화로 인한 누적자승 오차의 변화량이고, 율의 변화량 $\Delta R(e)$ 는 질감 및 율왜곡 화소의 부호화에 소요되는 비트 수의 변화량이다.

본 논문에서 율의 변화량은 실제로 부호화되는 비트 수를 조사할 경우의 복잡도가 대단히 큼을 감안하여 근사적으로 추정하기로 한다. 그리고, 질감에 소요되는 비트 수는 사용되는 질감 표현 기법에 관계없이 한 영역에 대하여 동일하게 할당되는 것으로 한다. 그러면, 임의의 가치 e 에서 발생하는 율의 변화량 $\Delta R(e)$ 은 다음 식으로 추정할 수 있다.

$$\Delta R(e) \approx - \{ a \times N_c(e) + \beta + \delta(e) \} \quad (10)$$

여기서 a 는 율왜곡화 소 당 비트 수, $N_c(e)$ 는 e 가치에 연결된 두 영역이 공유하는 율왜곡화 소 수, 그리고 β 는 한 영역의 질감을 표현하는 데 필요한 고정된 비트 수를 나타낸다. 또한, $\delta(e)$ 는 e 와 연관된 영역이 독립된 영역일 때 율왜곡의 새로운 시작점임을 나타내는 데 필요한 비트 수를 나타낸다.

나아가, 본 논문에서 각 영역의 질감은 평균값기호만 근사화 하는 것으로 하였다. 그러면, 영역 R_i 의 질감 근사화로 인한 왜곡은 다음과 같이 표현된다.

$$D(R_i) = \sum_{(x,y) \in R_i} [I(x,y) - \mu(R_i)]^2 \quad (11)$$

여기서 $I(x,y)$ 는 원 영상의 밝기이고, $\mu(R_i)$ 는 영역 R_i 의 평균값이다. 따라서, 임의의 가치 e 에 연결된 두 영역 R_i 와 R_j 를 병합할 때 발생하는 왜곡의 변화량 $\Delta D(e)$ 는 (11)식을 (2)식에 대입함으로써 다음과 같이 주어진다^[5].

$$\begin{aligned} \Delta D(e) &= D(R_i \cup R_j) - [D(R_i) + D(R_j)] \\ &= |R_i| \mu(R_i)^2 + |R_j| \mu(R_j)^2 - (|R_i| + |R_j|) \mu(R_i \cup R_j)^2 \quad (12) \\ &= (\mu(R_i) - \mu(R_j))^2 (|R_i|^{-1} + |R_j|^{-1})^{-1} \end{aligned}$$

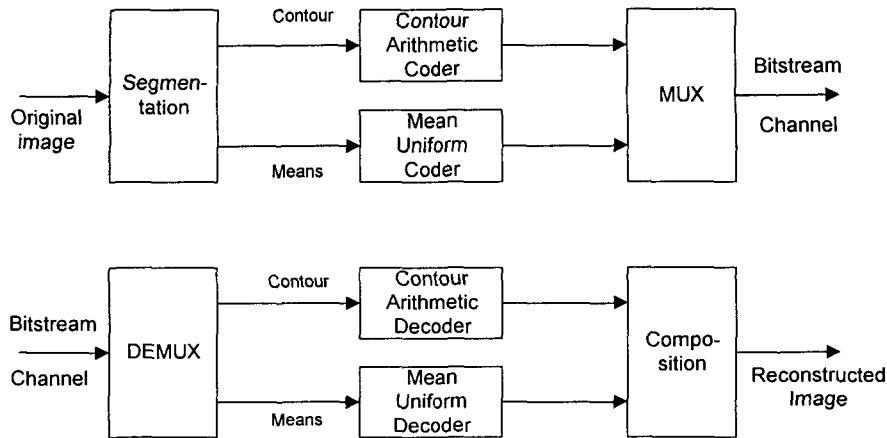


그림 6. 실험에 사용된 영역기반 부·복호화기의 블록도

Fig. 6. Block diagram of the region-based coder/decoder used in simulation.

여기서 $|R_i|$ 는 영역 R_i 의 크기(화소 수)이다.

4. 올웨곡 기반 반복적 병합

지금까지 기술된 올웨곡 기반 비용과 Lagrange 상수 추정법에 근거한 올웨곡 기반 반복적 병합법은 다음과 같이 수행된다. 우선, 초기 분할된 영상으로부터 초기 RAG를 구성한다. 이때 각 가지의 비용은 (6)식 및 (10)~(12)식에 근거하여 계산한다. 다음, 각 단계 t 에서 최소 비용 가지 $e_{opt} = \arg \min_{e \in E_t} \{\Delta J(e)\}$ 를 찾아서, 이 가지에 연결된 두 영역 R_i 와 R_j 를 서로 병합한다. 이때 병합된 영역은 새로운 영역 $R_k \cup R_l \in P_{t+1}$ 을 형성함으로써, $t+1$ 단계에서의 최적의 파티션 P_{t+1} 을 얻게 된다. 그리고, 새로운 영역과 연관된 가지들에 대하여 올 및 왜곡의 변화량을 갱신하여, P_{t+1} 에 대응하는 새로운 RAG를 구성한다. 이러한 과정을 그림 3과 같이 미리 정해진 영역 수 혹은 왜곡 수준에 도달할 때까지 반복한다.

5. 영역기반 부호화에 적용

본 논문에서는 이렇게 올웨곡 기반 반복적 병합에 의해서 영상 분할을 수행하고, 그 결과를 영역기반 부호화 하였다. 영역기반 부호화에서, 영역의 질감은 평균밝기로 표

현하여 고정길이 부호화하고, 영역들의 윤곽은 윤곽 상태를 이용한 여섯 방향 윤곽부호화^[12]에 의해서 발생된 심벌들을 산술부호화 하였다. 실험에 사용된 영역기반 부·복호화기는 그림 6과 같으며, 실제로 비트 스트림을 발생시키고 이들 비트 스트림으로부터 영상을 재생한다.

IV. 실험결과 및 고찰

본 논문에서는 제안한 올웨곡 기반 비용을 사용하여 반복적 병합하는 영상 분할법의 성능을 알아보기 위하여 256×256 크기의 8bit Lena 및 Pepper 영상에 대하여 실험하였다. 그림 7은 실험에 사용된 원 영상들이다.

제안한 올웨곡 기반 비용에서, ΔR 의 계산에 필요한 α , β 및 $\delta(e)$ 를 각각 $\alpha=1.4$ bits/pixel, $\beta=8$ bits/region 및 $\delta(e)=18$ bits로 설정하였다. 기존의 평균밝기 차이 ΔM , 왜곡의 변화량 ΔD , 올웨곡 기울기 $\Delta D/\Delta R$ 을 이용해서 분할한 결과들의 부호화 성능과도 비교하기 위하여, 그림 6의 영역기반 부·복호화기로 부·복호화를 수행하였다. 또한, 기존의 정지영상 부호화의 표준 기법인 JPEG 과도 부호화 성능을 비교하였다. 이때, 복원된 영상의 화질에 대한 객관적 척도로는 일반적으로 많이 사용하는 다음과 같은 PSNR을 화질평가 척도로 사용하였다.



그림 7. 실험에 사용된 원 영상들 (a) Lena, (b) Pepper
 Fig. 7. Original images used in simulation: (a) Lena and (b) Pepper

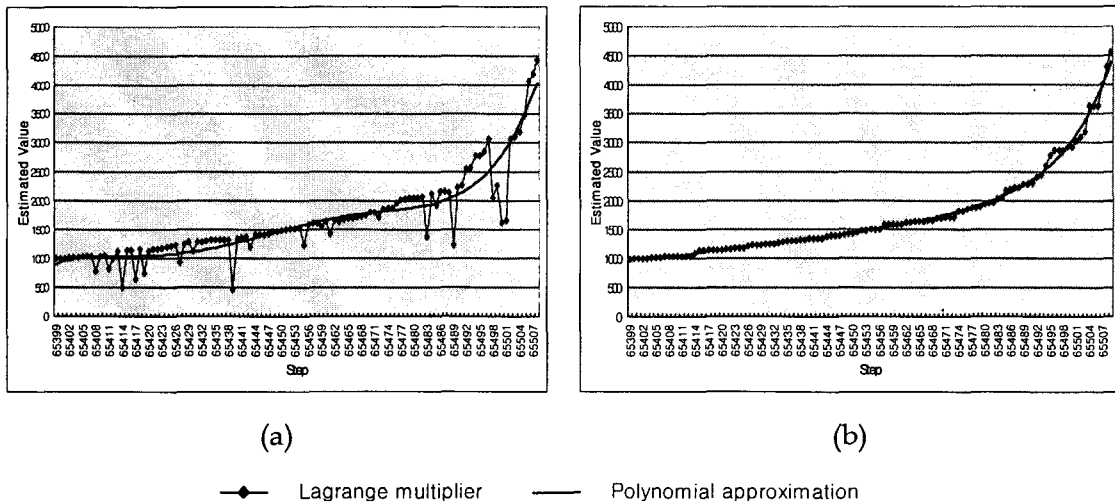


그림 8. 병합단계에 따른 추정된 Lagrange 상수 곡선: (a) 윌왜곡 기울기 기반, (b) 제안한 방법
 Fig. 8. Curve of estimated values of Lagrangian multiplier based on (a) rate-distortion slope and (b) the proposed method.

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{N^2} \sum_i \sum_j [K(i, j) - \hat{I}(i, j)]^2} \text{ [dB]} \quad (13)$$

여기서 $N \times N$ 은 영상 크기이고, $I(i, j)$, $\hat{I}(i, j)$ 는 원 영상과 재생영상의 밝기이다.

본 논문에서는 제안한 윌왜곡 기반 비용에 사용되는 Lagrange 상수를 (9)식으로 추정하여 영상 분할을 수행하

였다. 이때 λ_0 는 실험적으로 $\lambda_0 = 10.0$ 을 설정하였다. 그림 8은 병합단계에 따른 (7)식과 (9)식으로 추정된 Lagrange 상수 값들의 곡선을 각각 나타낸다. (7)식에 의한 결과인 (a)에서는 실제 윌왜곡의 기울기에 비해 매우 작은 값으로 추정되는 경우가 있으나, (9)식에 의한 (b)에서는 상수 값이 부드럽게 단조 증가해 감을 알 수 있다.



그림 9. Lena 영상에 대하여 반복적 병합법에 의한 동일한 PSNR 26.52 dB에서의 윤곽선 영상들: (a) 평균밝기 차이, (b) 왜곡의 변화량, (c) 울왜곡 기울기, (d) 제안한 비용에 의한 기법

Fig. 9. The contour maps at the same PSNR 26.52 dB for Lena image by recursive merging based on (a) mean-difference, (b) distortion, (c) rate-distortion slope, and (d) the proposed method.

표 1은 Lena 영상을 다양한 비용기반 분할법을 사용하여 분할한 후 영역기반 부호화를 수행했을 때의 성능을 조사한 것이다. 이때, 분할정도는 사용된 분할법에 관계없이 재생영상들이 동일한 PSNR 26.52 dB를 가지도록 조정되었다. 표에서 보면, 제안한 분할법을 기존의 평균밝기 차이에 의한 분할법과 비교했을 때, 영역 수와 윤곽화소 수가 대폭 줄어들었기 때문에 비트율이 0.26 bpp 정도 감소하였다. 그리고, 왜곡의 변화량에 의한 분할법과 비교했

을 때, 영역 수는 많이 늘어났지만 윤곽화소 수가 대폭 줄어들었기 때문에 비트율이 0.07 bpp 정도 감소하였다. 또한, 울왜곡 기울기에 의한 분할법과 비교했을 때에도, 영역 수는 다소 증가하지만 윤곽화소 수는 오히려 감소하여 전체적으로는 비트율이 0.01 bpp 정도 감소하였다. 이것은 제안한 분할법의 윤곽선 영상이 울왜곡 기울기에 의한 분할법에 의한 윤곽선 영상에 비해서 윤곽선이 다소 부드러워졌음을 의미한다.

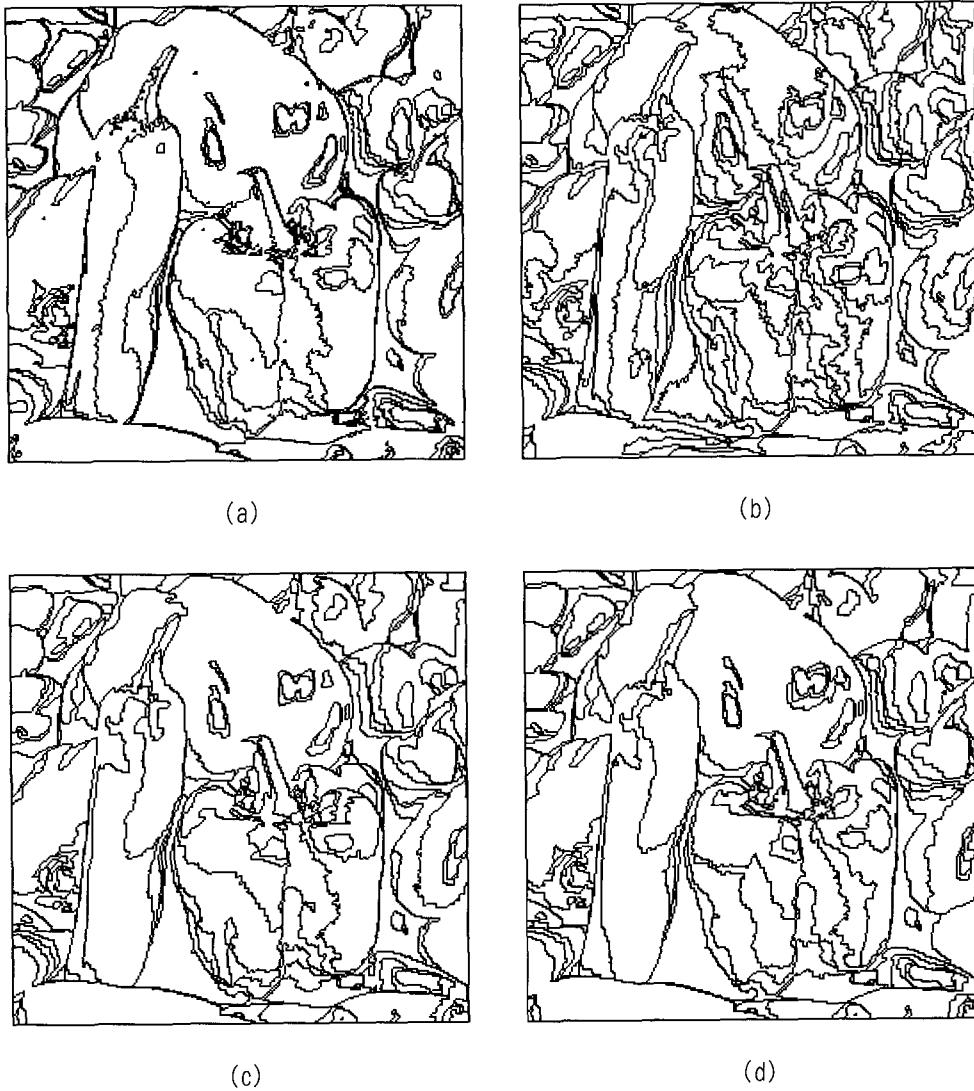


그림 10. Pepper 영상에 대하여 반복적 병합법에 의한 동일한 PSNR 27.51 dB에서의 윤곽선 영상들: (a) 평균값기 차이, (b) 왜곡의 변화량, (c) 율왜곡 기울기, (d) 제안한 비용에 의한 기법
 Fig. 10. The contour maps at the same PSNR 27.51 dB for Pepper image by recursive merging based on (a) mean-difference, (b) distortion, (c) rate-distortion slope, and (d) the proposed method.

표 2는 Pepper 영상을 다양한 비용기반 분할법을 사용하여 분할한 후 영역기반 부호화(PSNR=27.51 dB)를 수행했을 때의 성능을 조사한 것이다. 표 2의 결과는 표 1의 Lena 영상에 대한 결과에서와 같은 유사한 특성을 나타내고 있다. 표에서 보면, 제안한 분할법을 평균값기 차이와 왜곡의 변화량 및 율왜곡의 기울기에 의한 분할법과 비교하면, 비트율이 0.21, 0.08, 0.01 bpp 정도 감소하는 것으로

나타났다.

그림 9는 표 1의 Lena 영상에 대한 분할 결과들의 윤곽선 영상을 나타낸 것이다. 평균값기 차이에 의한 윤곽선 영상에는 입력영상에 존재하는 과립성 변화로 인하여 소영역이 많이 발생한다. 이러한 소영역들은 화질의 개선에는 도움을 주지 못하고, 윤곽 및 질감을 부호화 하는 데 많은 비트가 소요되므로 부호화 성능을 저하시키는 결정

표 1. Lena 영상에 대한 동일한 PSNR 26.52 dB에서의 각 비용에 따른 영역기반 부호화 성능
Table 1. Performance of region-based coding according to each cost for Lena image at the same PSNR 26.52 dB.

Cost	Bitrate[bpp]	Number of regions	Number of contour pixels
Mean difference	0.58	1,430	14,378
Distortion	0.39	182	16,495
Rate-distortion slope	0.33	250	13,054
Proposed method	0.32	259	12,835

표 2. Pepper 영상에 대한 동일한 PSNR 27.51 dB에서의 각 비용에 따른 영역기반 부호화 성능
Table 2. Performance of region-based coding according to each cost for Pepper image at the same PSNR 27.51 dB.

Cost	Bitrate[bpp]	Number of regions	Number of contour pixels
Mean difference	0.50	1,159	13,639
Distortion	0.37	176	15,500
Rate-distortion slope	0.30	260	11,984
Proposed method	0.29	267	11,774

적인 요인이 된다. 왜곡의 변화량에 의한 윤곽선 영상에는 다른 방식에 비하여 가장 작은 수의 영역들이 존재하고 소 영역은 거의 존재하지 않음을 알 수 있다. 그러나, 영역 수는 작은 대신 윤곽화소 수가 가장 크고 윤곽선의 모양이 다소 복잡하여, 윤곽을 부호화 하는 데 비트가 많이 소요된다. 이에 반하여, 율왜곡 기울기에 의한 분할법 및 제안한 분할법은 윤곽선을 고려하여 분할을 수행했기 때문에 영역 수가 많지 않으면서도 소 영역 및 윤곽화소 수가 상당히 줄어들었고 윤곽선도 단순해 졌음을 알 수 있다. 특히, (9)식에 의한 제안한 방법의 윤곽선이 다소 부드러워 보인다. 그림 10은 표 2의 Pepper 영상에 대한 분할 결과들의 윤곽선 영상을 나타낸 것이다. 그림 10은 그

림 9의 Lena 영상에 대한 윤곽선 영상들의 결과에서와 같은 유사한 특성을 나타내고 있다.

표 3은 Lena 영상을 다양한 비용기반 분할법에 대하여 동일한 비트율 0.40 bpp로 영역기반 부호화를 수행했을 때의 성능을 조사한 것이다. 표에서 보면, 제안한 분할법을 평균값기 차이 및 왜곡의 변화량에 의한 분할법과 비교했을 때, 재생영상의 PSNR이 각각 3.26, 0.89 dB 증가한다. 그리고, 율왜곡 기울기에 의한 분할법과 비교했을 때, 영역 수와 윤곽화소 수가 다소 증가했기 때문에 PSNR이 0.10 dB 증가한다. 표 4는 Pepper 영상에 대하여 동일한 비트율 0.37 bpp로 영역기반 부호화를 수행했을 때의 성능을 조사한 것이다. 표 4의 결과는 Lena 영상에

표 3. Lena 영상에 대한 동일한 비트율 0.40 bpp에서의 각 비용에 따른 영역기반 부호화 성능
 Table 3. Performance of region-based coding according to each cost for Lena image at the same bitrate 0.40 bpp.

Cost	PSNR [dB]	Number of regions	Number of contour pixels
Mean difference	24.22	839	10,590
Distortion	26.59	189	16,780
Rate-distortion slope	27.38	403	14,912
Proposed method	27.48	430	14,957

표 4. Pepper 영상에 대한 동일한 비트율 0.37 bpp에서의 각 비용함수에 따른 영역기반 부호화 성능
 Table 4. Performance of region-based coding according to each cost for Pepper image at the same bitrate 0.37 bpp.

Cost	PSNR [dB]	Number of regions	Number of contour pixels
Mean difference	25.24	735	10,768
Distortion	27.57	180	15,545
Rate-distortion slope	28.60	400	14,027
Proposed method	28.75	405	14,132

대한 결과에서와 같은 유사한 특성을 나타내고 있다. 표를 보면, 제안한 분할법을 평균밝기 차이와 왜곡의 변화량 및 울왜곡의 기울기에 의한 분할법과 비교하면, PSNR이 3.51, 1.18, 0.15 dB 증가하는 것으로 나타났다.

그림 11은 표 3의 Lena 영상에 대한 분할 결과들의 재생영상을 나타낸 것으로, 제안한 비용과 울왜곡 기울기에 의한 분할법은 기존의 평균밝기 차이 및 왜곡의 변화량에 의한 분할법들에 비해 PSNR 관점에서나 주관적 관점에서 우수해 보인다. 그러나, 본 실험에서 사용한 영역기반 부호화에 의해서 영역의 내부 질감을 영역의 평균밝기로만 표현했기 때문에 영역간의 경계에서 급격한 밝기 변화

를 보이고 있는 것을 볼 수 있다. 그러나, 이러한 현상은 후처리 과정을 거치거나 질감을 표현하는 기법을 달리 한다면 제거 가능하다.

그림 12는 Lena 영상에 대하여 제안한 분할법, 기존의 분할법들 및 JPEG의 부호화 성능을 부호화 비트율에 따른 PSNR 성능 곡선을 나타낸 것이다. 그림에서, 제안한 분할법은 PSNR 측면에서, 평균밝기 차이 및 왜곡의 변화량을 이용한 분할법에 비하여 약 2.5~3.5 dB, 0.8~1.0 dB 정도의 성능 개선이 있었고, JPEG 보다는도 0.3~0.6 dB 성능 개선이 있었다.

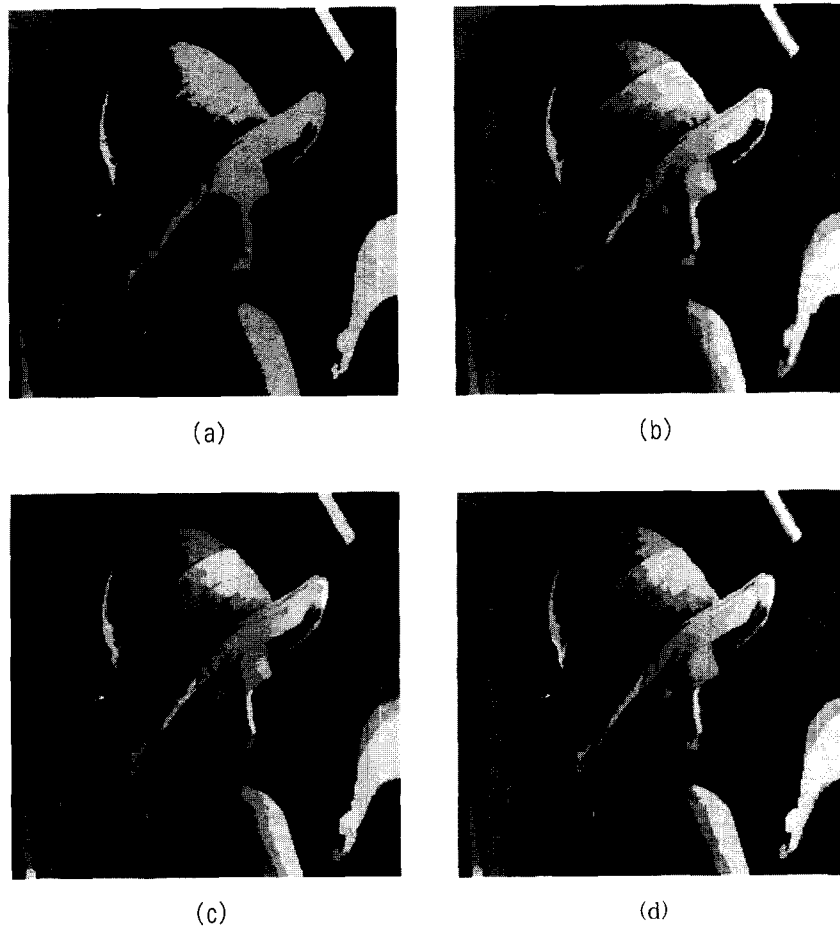


그림 11. Lena 영상에 대하여 반복적 병합법에 의한 동일한 비트율 0.40 bpp에서의 재생 영상들: (a) 평균차이 차이, (b) 왜곡의 변화량, (c) 울왜곡 기울기, (d) 제안한 비용에 의한 가법
 Fig. 11. The reconstructed images at the same bitrate 0.40 bpp for Lena image by recursive merging: (a) mean-difference, (b) distortion, (c) rate-distortion slope, and (d) the proposed method.

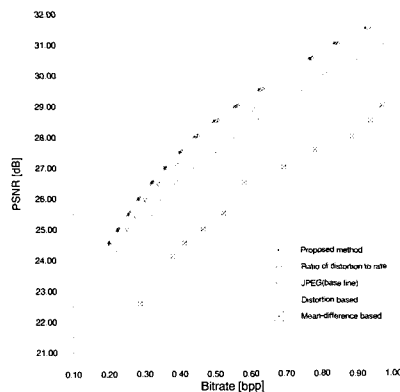


그림 12. Lena 영상에 대한 영역기반 부호화와 JPEG과의 PSNR 성능 비교
 Fig. 12. Comparison of PSNR performance of region-based coding and JPEG for Lena.

V. 결 론

참 고 문 헌

본 논문에서는 영역기반 분할법의 일종으로 RAG를 이용한 반복적 병합법에 의한 울왜곡 기반 영상 분할을 제안하였다. 이때, 인접한 두 영역간의 이질성의 비용으로 울왜곡 기반 비용을 정의하고, 이 비용에 근거하여 영상 분할을 수행하였다. 그리고, 이러한 울왜곡 기반 비용에 의한 분할법의 부호화 성능을 기존의 평균밝기 차이, 왜곡의 변화량 그리고 울왜곡 기울기를 이용한 비용들에 의한 분할법들의 부호화 성능들을 함께 비교하여 분석하였다. 이때 사용된 영역기반 부호화법은, 영역의 질감은 평균밝기로 표현하여 고정길이를 부호화하고, 영역들의 윤곽은 여섯 방향 윤곽부호화법에 의하여 윤곽화소들을 따라가면서 부호화하였다.

제안한 비용은 각 병합 단계마다 울왜곡 관점에서 최적의 조건을 만족하는 두 영역을 병합하기 위하여, 울의 변화량과 왜곡의 변화량을 동시에 고려하도록 정의하였다. 이때, 사용되는 Lagrange 상수는 현재 단계에서의 울왜곡 기울기의 최소값과 이전 단계에서의 추정된 Lagrange 상수 및 초기 설정값을 같이 고려하여 추정하였다.

본 연구를 통하여, 울의 변화량과 왜곡의 변화량을 함께 고려한 울왜곡 기반 비용을 제시하고, 반복적 병합에 의한 울왜곡 기반 영상 분할의 이론적 근거를 마련하였다. 그리고, 기존의 울왜곡 기울기를 이용한 분할법이 울왜곡 기반 비용을 사용한 분할법의 특수한 경우임을 보였다. 실험 결과, 평균밝기 차이에 의한 분할법은 소 영역이 많이 발생하고, 왜곡의 변화량에 의한 분할법은 윤곽화소 수가 크게 증가하고 윤곽선이 복잡하다는 단점이 있음을 알 수 있었다. 이에 반하여, 울왜곡 기반 분할법은 윤곽화소 수가 적고 윤곽선의 특성이 단순하여 기존의 평균밝기 차이, 왜곡의 변화량 그리고 울왜곡 기울기를 이용한 비용들에 의한 분할법에 비하여 동일한 비트율에서 2.5~3.5 dB, 0.8~1.0 dB, 0.1~0.2 dB 정도의 PSNR 성능 개선이 있었다. 또한, 정지영상 부호화 표준 기법인 JPEG과 비교하여 0.3~0.6 dB 정도의 PSNR 성능 개선이 있음을 알 수 있었다.

앞으로, Lagrange 상수를 보다 정확하게 추정하여 부호화 성능을 개선시키는 연구가 계속되어야 하겠고, 분할 영상의 영역 내부의 질감 표현 기법을 개선하여 보다 자연스럽게 영상을 재생하는 것에 관한 연구가 추가되어야 하겠다. 그리고, 분할 수행 시간을 단축하기 위한 노력도 있어야 하겠다.

[1] R. C. Gonzalez and R. E. Wods, *Digital Image Processing*, Addison Wesley, 1993.

[2] A. K. Jain, *Fundamental of Digital Image Processing*, Printice-Hall, Englewood Cliffs, NJ, 1989.

[3] S. L. Horwitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *Journal of ACM*, vol. 23, no. 2, pp. 368-388, Apr. 1976.

[4] ISO/IEC JTC1/SC29/WG11, N1992, MPEG-4 Video Verification Model Version 11.0, Mar. 1998.

[5] M. J. Biggar, O. J. Morris, and A. G. Constantinides, "Segmented-image coding: A performance com-parison with discrete cosine transform," *IEE Proc.*, vol. 135, pt. F, pp. 121-132, Apr. 1988.

[6] M. Kunt, M. Benard, and R. Leonardi, "Recent results in high-compression image coding," *IEEE Trans. Circuits and Sysytems*, vol. CAS-34, no. 11, pp. 1306-1336, Nov. 1987.

[7] R. M. Harlalick and L. G. Shapiro, "Survey : Image segmentation technique," *CVGIP*, vol. 29, pp. 100-132, 1985.

[8] M. Kocher and R. Leonardi, "Adaptive region growing technique using polynomial functions for image approximation," *Signal Processing*, vol. 11, pp. 47-60, 1986.

[9] O. J. Morris, M. de J. Lee, and A. G. Constantinides, "Graph theory for image analysis: An approach based on the shortest spanning tree," *IEE Proc.*, vol. 133, pt. F, no. 2, pp. 146-152, Apr. 1986.

[10] S. H. Kwok and A. G. Constantinides, "A fast recursive shortest spanning tree for image segmentation and edge detection," *IEEE Trans. on Image Processing*, vol. 6, no. 2, pp. 328-332, Feb. 1997.

[11] H. Sanderson and G. Crebbin, "Image segmentation for compression of image and image sequences," *IEE Proc. Vis. Image Signal Process.*, vol. 142, no. 1, Feb. 1995.

[12] 홍원학, 허진우, 김남철, "윤곽 상태를 이용한 여섯

방향 윤곽부호화”, 방송공학회 논문지, 제 1권 제 1호, 1996년 6월.

Rate-Distortion Based Video Compression, Kluwer Academic Publishers, 1997.

[13] G. M. Schuster and A. K. Katsaggelos,

저 자 소 개



전 성 철

1971년 3월 28일생.
 1997년 2월 경북대학교 전자공학과 졸업(공학사).
 1999년 2월 경북대학교 대학원 전자공학과 졸업(공학석사).
 1999년 ~ 현재 현대전자 연구원.
 주관심분야 : 영상신호처리, 영상압축.



임 채 환

1971년 1월 5일생.
 1993년 2월 경북대학교 전자공학과 졸업(공학사).
 1995년 2월 경북대학교 대학원 전자공학과 졸업(공학석사).
 1995년 6월~현재 경북대학교 대학원 전자공학과 박사과정.
 주관심분야 : 영상신호처리, 컴퓨터 비전, 영상압축.



김 남 철

1955년 7월 29일생.
 1978년 2월 서울대학교 전자공학과 졸업(공학사).
 1980년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사).
 1984년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학박사).
 1984년 3월~현재 : 경북대학교 전자·전기공학부 교수.
 1991년 8월~1992년 8월 미국 Syracuse 전기 및 전산학과 객원교수.
 1990년 1월~1994년 12월 생산기술연구원 HDTV 신호처리분과위원.
 1994년 1월~1996년 12월 한국통신학회 편집위원.
 1994년 1월~1996년 12월 대한전자공학회 편집위원.
 1996년 1월~현재 한국방송공학회 편집위원, 이사.
 1996년 11월~현재 국방과학연구소 국방과학기술분류 및 수준 조사위.
 1997년 1월~1998년 12월 한국통신학회 신호처리 연구회 위원장.
 주관심분야 : 영상신호처리, 영상압축, 영상복원, 컴퓨터 비전.