

내용기반의 MPEG-4 객체 추출 연구

Content-based MPEG-4 Object Extraction

권기훈, 최석림

세종대학교 컴퓨터공학과, 전자공학과

요 약

본 논문에서는 연속적인 입력화상에서 움직임을 나타내는 객체(Object)를 적은 연산량을 사용하여 추출해 내는 알고리즘을 소개한다. 본 알고리즘은 두 가지 단계로 이루어진다. 첫번째 단계로, 이전의 영상과 현재의 영상을 비교하여 움직임의 변화를 보이는 영역을 찾는다. 이 단계에서는 영상을 비교하여 움직임을 추출하기 위하여 참조영상과 현재의 영상, 그리고 영상의 데이터로서 edge 정보를 사용한다. 두 번째 단계에서는, 첫번째 단계에서 움직임으로 판단된 Object mask(변화를 가지는 영역)를 가지고 background 제거 및 Object의 정확한 shape을 만들기 위한 post-processing 과정을 가지게 된다. 이 두 단계를 거친 후 입력영상에서 background를 떼어낸 최종적인 Object의 shape 정보가 추출되게 된다. 이 알고리즘은 object를 기반으로 부호화함으로써 데이터의 압축률을 극대화 시키는 MPEG-4 뿐만 아니라, video database, 무선 통신등과 같은 다양한 범위의 application에 적절하게 사용될 수 있을 것이다.

1. 소 개

요구되는 영상의 질을 보장하면서 낮은 전송률을 가지는 데이터를 생성하기 위하여 block-based hybrid coding 기법을 사용하는 기존의 MPEG 표준과는 달리, 현재 표준화과정이 진행중인 MPEG-4에서는 content-based coding 방식을 바탕으로 한다. 즉, 영상 내에서 의미를 가지는 객체(Object)를 바탕으로 coding 하는 개념으로서, 이는 다음과 같은 여러 가지의 기능들을 가능하게 한다. 우선, 부호기측에서는 객체를 바탕으로 선택적인 부호화를 가능하게 함으로써 계산량을 줄이는 동시에 주어진 대역폭(bandwidth)에 적절하게 대처할 수 있으며 수신측에서는 복호기의 사양에 맞게 객체를 바탕으로 한 선택적인 복호화를 함으로써 최소한의 데이터로 실시간 디스플레이를 가능하게 한다. 또한, 사용자의 입력에 의해 얻어진 각 객체간 재구성(recomposition)을 가능하게 함으로써 사용자가 직접 원하는 동영상을 구현할 수 있는 user interaction을 가능하게 한다. 본 논문은 낮은 complexity를 가지는 부호기에서 사용될 수 있는 MPEG-4에 정의되어 있는 연속된 영상내의 움직

임을 가지는 객체(Object)를 추출하는 기법을 제안하였다. 움직임을 갖는 Object의 shape을 추출하기 위해서 기존에 사용되어온 기법들은 이전의 영상과 현재의 영상의 조도값을 비교하는 방법을 사용하여 왔다. 이러한 기법들은 사용하는 영상의 데이터에 따라 크게 두 가지로 나뉘어 질 수 있다. 하나는 영상의 데이터로서 luminance 정보를 사용하는 방법이 있고, 나머지 하나는 edge 정보를 사용하는 방법이다. 영상의 luminance 차이를 사용한 기법은 조도에 상당히 민감한 결과를 나타내며, 두 영상간의 luminance 차에 대한 threshold를 적용하는데 다소 어려운 점이 있다. 반면, Edge 정보를 사용한 기법은 다음의 장점을 가진다. 첫째, edge는 조도의 변화에 민감한 반응을 보이지 않는다는 것이다. 이것은 움직임 탐지를 위해 두 영상을 비교하였을 때 외부환경의 변화나 영상신호 자체의 잡음에 따라서 발생할 수 있는 noise에 큰 영향을 받지 않는다는 것을 말한다. 둘째, edge 정보를 사용함으로써 영상내의 많은 양의 데이터를 줄일수 있다는 장점을 가질 수 있다. 이것은 제한된 processing power를 갖는 hardware에 있어서 효과적이며 coding 효율면에서 잇점을 가질 수 있을 것이다. 본 논문에서는 noise robust한 segmentation과 shape추출을 위해 edge정보와 luminance정보를 각 module에 적절히 적용함으로써 정확한 motion object를 추출하는 기법을 보이도록 하겠다. 알고리즘은 두 가지의 단계를 가진다. 첫번째 단계는 bi-level edge를 사용하여 motion을 탐지하는 부분이다. 이 단계에서는 영상내에서 filtering을 통하여 추출된 gray level edge값을 bi-level edge로 바꾸기 위한 adaptive threshold가 적용된다. 두번째 단계는, 첫번째 단계에서 생성된 motion mask내에서 background에 해당되는 부분을 제거하고 정확한 object shape을 생성해 내는 단계이다

2. The segmentation algorithm

이 단계에서는 참조영상과 현재영상을 비교하여 motion object를 포함하는 직사각형 모양을 갖는 Motion Object Mask(MOM)부분과 영상내에서 mask를 제외한 background부분을 분리해내는 것을 목적으로 한다. 이 단계에서 생성된 직사각형 모양의 MOM은 이후의 post-processing 단계를 거쳐 정확한 object shape을 보이게 된다. MOM추출을 위한 block diagram은 그림 1에서 보여진다.

2.1. Gray level edge의 bi-level edge 변환을 위한 adaptive threshold 값 추출

본 알고리즘은 bi-level edge 값을 입력으로 하여 motion object를 탐지한다. Edge filtering의 결과 발생한 Gray level edge는 적절한 thresholding에 의해 binary 값으로 변환되어야 한다. Threshold에 의해 변환된 binary edge 값은 입력영상의 밝기에 따라 그 결과가 다르게 나타남을 알 수 있다. 신뢰성 있는 edge 값을 얻기 위해서는 영상의 밝기에 대응해 적절한 threshold 값이 적용되어야 한다. Adaptive threshold와 bi-level edge 값은 다음식으로부터 산출되어질 수 있다.

$$\begin{cases} B(i) = 1 & \text{If } E(i)^2 > T(k - 2(x_size + y_size)) \\ B(i) = 0 & \text{else} \end{cases}$$

$$T = \sum_{i=0}^k E(i)$$

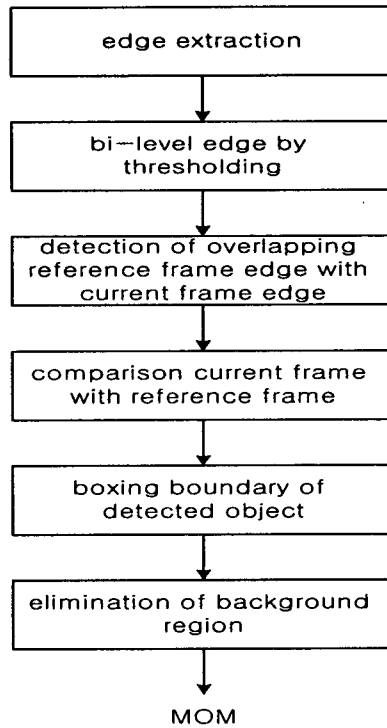


그림 1. Block diagram of MOM extraction

여기에서, E 는 luminance edge로 표현된 영상, B 는 bi-level edge에 대한 영상이며, k 는 전체 영상의 pixel 개수, x_size , y_size 는 영상의 수직, 수평 scan line에 대한 pixel의 개수를 나타낸다. 그림 2는 고정 threshold와 adaptive threshold를 적용했을 때 binary image의 결과를 보여준다.

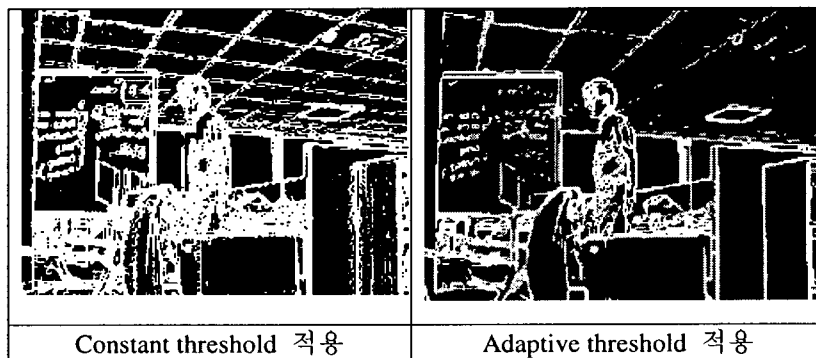


그림 2. Constant threshold와 adaptive threshold 적용에 의한 bi-level image 결과

2.2. 두 영상의 비교를 통한 motion 탐지

bi-level edge 값을 가지는 두 영상에 대해 pixel 단위의 logic operation을 사용하여 변화를 일으킨 부분을 탐지한다.

2.3. Edge-occlusion check

위의 bi-level 를 이용한 motion 탐지는 계산의 간단함과 edge 의 특징인 noise 에 민감하지 않는 장점이 있는 반면, 한가지 단점을 가지게 된다. 그것은 적용 operation 특성상 참조영상의 배경 edge 와 현재영상의 object 의 edge 가 겹쳐지는 경우 이 부분에서는 motion 이 탐지되지 못한다는 것이다. 이 문제를 해결하기 위하여 겹쳐진 pixel 에 대해서는 참조영상과 현재 영상간의 luminance 의 비교를 수행하게 한다. 이와 같이 motion 탐지 알고리즘은 다음과 같이 표현되어진다.

$$\begin{cases} \text{If } I_E(i,j) = 1 \text{ and } R_E(i,j) = 1 \\ \quad \text{If } I_L(i,j) - (R_L(i,j)) > th \text{ then, } O(i,j) = 1 \\ \text{Else } O(i,j) = I_E(i,j) \ \&\& \ \sim(R_E(i,j)) \end{cases}$$

여기에서, i, j 는 영상내의 pixel 의 위치를 나타내며 I_E, I_L 은 각각 edge 값과 luminance 값을 가지는 입력영상을 나타내며 R_E, R_L 은 edge 값과 luminance 값을 가지는 참조영상, O 는 두 영상의 비교에 대한 결과를 가지는 영상이다. 그림 3 은 motion object 탐지에 대한 결과를 보여준다.

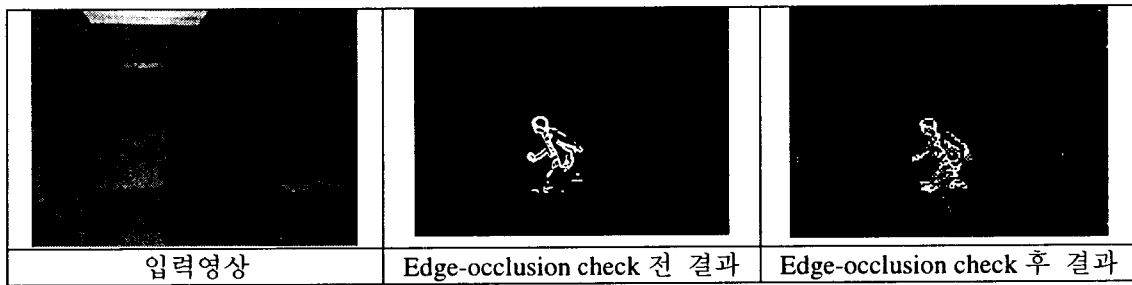


그림 3. Edge-occlusion 처리의 비교 및 Motion object 탐지 결과

2.4. Motion Object Mask (MOM)

위의 단계를 거친 결과는 단지 변화된 pixel 을 판단하여 motion 을 가지는 부분을 추출한 것이므로 정확한 object shape 추출을 위한 처리가 필요할 것이다. MOM 은 최종적으로 전체 영상에서 motion object 에 해당되는 부분에 bounding box 를 하고 그 외의 background 영역을 제거한 것으로써, 이것은 이후의 object shape 의 생성을 위한 Post-processing 과정에서의 불필요한 계산량을 감소시키는 효과를 준다. 그림 4 는 motion 에 해당되는 부분을 추출한 후 background 영역을 제거한 Motion Object Mask 를 나타낸다.

3. Post-processing

이 과정은 1 차적으로 움직임으로 판단된 object mask(MOM)내에서 object shape 을 가지는 객체를

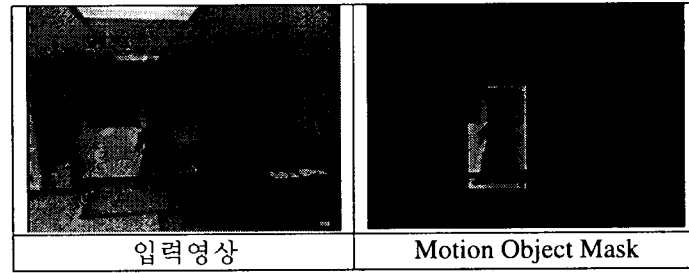


그림 4. Motion Object Mask(MOM)의 결과

생성하는 과정이다. Post-processing 과정은 두 가지 단계를 거쳐 최종적인 object shape 을 추출하게 된다. 첫번째 단계는 MOM 내에서 background 영역을 제거하는 부분이다. 이것은 이전 영상의 object shape 에 해당하는 영역과 현재의 MOM 마스크를 사용하여 이루어진다. 이 때 displacement vector information 를 사용하여 uncovered background 영역에 해당하는 부분과 실제 object 에 해당하는 부분을 판단하여 background 영역을 제거한다. Displacement 의 판단은 hierarchical block matching 기법을 사용하여 이루어진다. 첫번째 단계의 결과 후 edge filtering 과정을 거쳐 object 의 boundary 를 추출함으로써 정확한 object 정보를 생성할 수가 있다.

4. 실험

소개된 motion object 추출 알고리즘은 360 x 240 bmp file format 을 가지는 test video sequence 상에서 수행되었다. 여기에서 카메라의 움직임은 없다고 전제한다. 알고리즘의 적용결과는 그림 5 에서 보여지는 것과 같이 motion estimation 후의 MOM 상에서 background 를 제거 후 edge 추출 과정을 거쳐 최종적인 motion object 에 해당하는 부분의 shape 이 추출되어졌다.

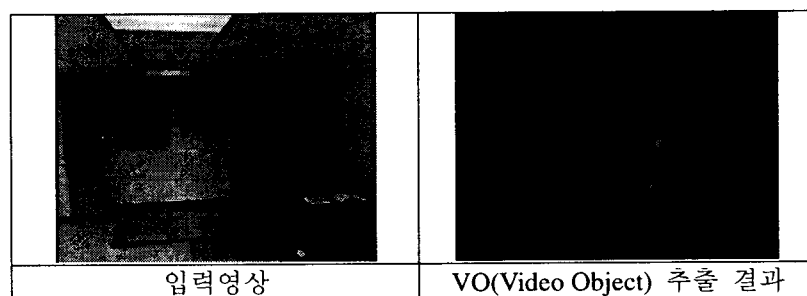


그림 5. VO (Video Object) segmentation

5. 결론

이 논문에서는 low processing power 를 가지고 video sequence 상에서 움직임을 가지는 object 를 추출하는 알고리즘을 소개하였다. 이 알고리즘은 크게 두 가지 단계를 거쳐 진행되었다. 첫 단계로는, bi-level edge 정보와 luminance 정보를 사용하여 motion object 에 해당하는 부분(MOM)을 추출하는 단계이다. 여기에서 motion 이 발생한 부분을 탐지하기 위하여 edge 정보와 luminance 정보를 사용하였다. 이것은 noise 에 robust 한 특성을 가지는 bi-level edge 정보를 사용, 간단한 연산을 통하여 motion 을 탐지하는데 따른 단점을 보완하기 위해 luminance 정보를 적용하는 이유이다. 두 번째 단계로는 현재영상의 MOM 과 이전 영상의 object 정보를 사용하여 object shape 정보를 생성해내는 단계이다. 이 단계에서는 displacement 정보를 사용하여 MOM 내의 background 부분을 제거한다. 본 논문에서 제안된 기법은 연속적인 영상에서 object 에 해당되는 부분을 추출하는 내용으로서, 이는 object-based coding 개념을 바탕으로 하는 MPEG-4 에서 의미를 가지는 객체를 추출하는 과정에 사용될 수 있으며 video databases, wireless virtual reality interfaces, low-bandwidth video compression, security monitoring 등의 다양한 범위의 application 에서 object 의 추출 및 tracking 에 적절하게 사용될 수 있을 것이다.

6. 참고문헌

- [1] MPEG-4, "MPEG-4 overview", ISO/IEC JTC1 / SC29 /WG11 N2459, Atlantic, October 1998.
- [2] MPEG-4, "Information Technology – Coding of Audio-Visual Objects: Visual Committee Draft", ISO/IEC JTC1/SC29/WG11 N2202, Tokyo, March 1998.
- [3] Aleksej Makarov, Jean_Marc Vesin, Florian Reymond, "Intrusion detection robust to slow and abrupt lighting changes". Real-time Imaging, SPIE, Vol. 2661, pp.44-54, February 1996.
- [4] Phillip M.Ngan. "Motion Detection in Temporal Clutter" Industrial Research Limited Report 693. March 1997.
- [5] Roland Mech, Michael Wollborn. "A Noise robust method for segmentation of moving objects in video sequences", Proceedings of IEEE ICASSP97, pp.2657-60, 1997.
- [6] M.Bierling, "Displacement estimation by hierarchical blockmatching", 3rd SPIE Symposium on Visual Communications and Image Processing, Cambridge, USA, pp.942-951, November 1988.
- [7] Christopher Richard Wren, Ali Azarbayejani, Trevor Darrell, Alex Paul Pentland, "Pfinder:Real-time Tracking of the Human Body", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No.7, July 1997