

DTV방송에 대비한 H/W중심의 MPEG Bitstream에서의 실시간 장면변환 검출방법

장 경 훈, 이 동 호

한양대학교 전자전기컴퓨터공학부

Real time Shot Change Detection in focus of H/W prepare for DTV broadcasting

Kyoung Hoon Jang, Dong Ho Lee

Department of Electrical & Computer Engineering, Hanyang University

khjang@image.hanyang.ac.kr

요 약

본 논문에서는 영상검색 기법에서 핵심인 Shot Change Detection 과 Non Linear Browsing 을 H/W 기반으로 구현하여, S/W 적으로는 비실시간으로만 가능하였던 video indexing 을 DTV 에 적용하여 실시간으로 구현하는 방법을 제시한다. 이를 위해 H/W part 는 실시간으로 들어오는 방송용 MPEG-2 bitstream 을 full decoding 이 아닌 최소화된 VLD (Variable Length Decoding) 수준의 parsing 으로 picture 내의 luminance 와 chrominance 의 DC 값, macroblock type, motion vector 정보를 얻어내어 각각의 histogram 을 계산하여 memory interface 를 통해 S/W 측에 넘겨주게 되고 S/W 는 각 상황에 맞게 indexing algorithm 을 변화시키며 최적의 video indexing 방법으로 확장할 수 있도록 하였다.

I. 서 론

국내에서도 시험방송이 시작된 DTV 에서 중요한 issue 중의 하나는 단지 MPEG-2 규격의 동영상을 시청하는 것 뿐 아니라 실시간으로 저장하여 사용자가 원하는 shot 의 검색을 할 수 있게 하고, 또한 non linear browsing 을 통해 검색속도를 조정하게 할 수 있게 하여 보다 편리한 검색을 가능케 하기 위한 방법을 찾는 것이다. 이미 S/W 적인 방법만으로는 여러가지 indexing 방법이 나와있으나 대부분은 MPEG bit stream 의 parsing 에 많은 load 가 소요되어 실시간으로 동영상에 검색 tag 를 첨가하는 것은 아직 실현되기 어렵다.[5] 게다가 High level 급의 방송용 bit stream 의 parsing 에는 더 많은 load 가 걸릴 것이다. 이러한

문제점을 해결하기 위해 본 논문에서 제안하고자 하는 것은 H/W 적으로 Shot Change Detection 이나 Non Linear Browsing 에 필요한 picture 내의 video indexing 에 관계된 parameter 만을 parsing 하는 것이다. Picture 내의 luminance 나 chrominance 의 변화는 shot change 를 찾아내게 할 수 있게 하고, motion vector 의 움직임 거리나 분포를 통해 non linear browsing 이 가능하게 된다. 하지만 단지 H/W 적인 parsing 만으로는 picture 내의 fade, dissolve, wipe 등을 포함한 shot change 의 detection 에는 엄연한 한계가 존재하기 때문에 이러한 정보를 단순 비교하는 것이 아니라 일정한 histogram 형태로 변환시켜 정보의 상관관계를 여러 가지 방법으로 비교하여 shot 의 촬영유형(shot type), 카메라 움직임(camera motion), 시선(gaze direction), 상의 위치(screen position)의 변환을 찾아낼 수 있는 algorithm 을 적용되어지도록 하고 S/W 적인 upgrade 를 통해 확장이 가능하도록 한다.

II. 전체 H/W 의 구조

MPEG 규격은 비디오의 복호에 대해서만 규정하고 있기 때문에 부호화기는 MPEG-2 규격에 정해져 있는 모든 mode 들을 지원해 줄 필요는 없으나, 복호기는 profile 과 level 에 해당되는 모든 mode 를 만족해야 하기 때문에 구현 및 검증에 있어 어려움이 있으며, 압축된 bitstream 내에 오류가 발생할 경우에 대한 대책을 마련해야 하며 고화질 영상의 복호를 위해서는 동작속도의 향상이 필수적이다. 이에 대한 모든 준비를 하려면 많은 H/W 적 load 와 logic 이 필요하게 되므로 picture 내의 특정 정보만을 얻어낼 수 있는 최소한의 logic 의

구현으로만 설계를 하였다. 전체 H/W 구조는 그림 1 과 같이 Pre_VLD, Decoding, Histogram block 의 세 부분으로 나뉘어 진다.

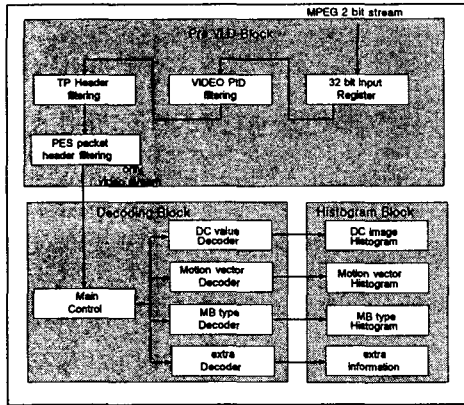


그림1. H/W 구조도

II-1. Pre VLD Block

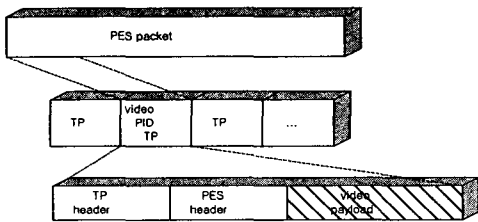


그림2. TP로부터 video stream의 추출

실제 방송용 TP 중에 있는 video TP 뿐만 아니라 audio, data, null TP 등이 들어오는데 이중 video TP 만의 PID 를 검사하여 filtering 하고 TP, PES packet 등의 header 와 adaptation field 또한 filtering 하며 shift register 를 이용 MPEG decoding 의 최대 단위인 32 bit video bitstream 을 만들어 주는 역할을 한다.

II-2. Decoding Block

Decoding 의 목적이 display 가 아니고 실시간 parsing 이기 때문에 buffer 를 이용한 buffer control 은 하지 않는다. MPEG stream 의 복호가 가능하도록 register 는 27MHz 한 clock 마다 한 bit 씩 shift 시키며 최대 코드 길이인 32bit 를 유지하며 decoding 된 bit 만큼 shift count 를 주어 항상 다음 부호를 복호 할 수 있는 상태를 유지토록 한다. 또한 각 picture 마다 사용되는 정보가 다르기 때문에 slice 단의 coding 부분은 logic 의 최소화와 speed 의 향상을

위하여 각 picture 별로 병렬로 동작된다. I-picture 에서는 DC value 를 끌어내기 위하여 DC code 와 AC code 를 해석하기 위한 부분을 위주로 설계되었고 P-picture 에서는 motion vector 추출을 위하여 macroblock 정보를 parsing 하는 부분을 위주로 설계하였다.

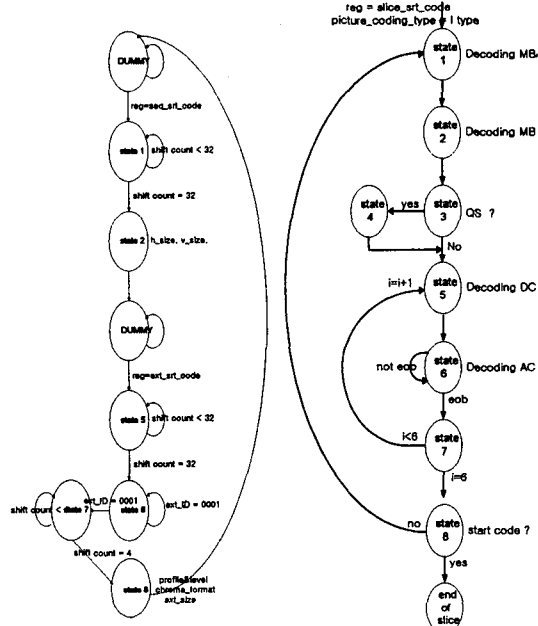


그림3. sequence state machine 그림4. I-picture state machine

II-3 Histogram Block

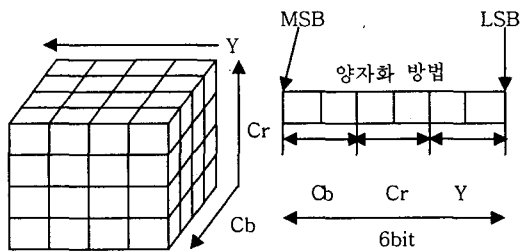


그림5. 64개의 bin으로 이루어진 DC image

DC image histogram 은 0~255 까지의 DC value Y, Cr, Cb 값을 각각 4 level 로 나누어 [그림 5]와 같이 64 개의 bin 으로 나누어 각 macroblock 별로 한 개씩의 bin counting 을 하게 된다. Y, Cr, Cb 를

따로따로 histogram 을 만들지 않고 세 값을 연계하여 하나의 histogram 을 만듦으로써 더 적은 sample 로 더 많은 종류의 색의 변화에 민감하게 대응할 수 있으므로 shot change detection 의 성능을 높일 수 있다. Shot change detection 의 방법은 이전 picture 와 현재 picture 와의 histogram 의 차가 임의의 threshold 값을 넘는다면 shot change 로 간주하고 다음 식으로 표현될 수 있다.

$$Difference = \sum_{k=1}^{NB} |Hist(P_k) - Hist(C_k)|$$

NB : number of bins(64)

Hist(P_k) : k-th bin of color histogram for previous picture

Hist(C_k) : k-th bin of color histogram for current picture

Motion vector histogram 은 P-picture 에서만 추출하며 motion vector 의 X,Y 좌표 값의 최대값은 f_code 에 의해 결정되므로 parsing 된 X,Y 좌표 값을 f_code 로 나누어 항상 같은 비율의 수평/수직 8 level 을 가지는 총 64 bin 의 histogram 을 만든다. 이 histogram 의 방향성과 크기에 따른 변화를 이용하여 움직임이 많은 영상과 움직임이 적은 영상 등을 구별해 내어 non linear browsing 에 응용할 수 있게 한다.

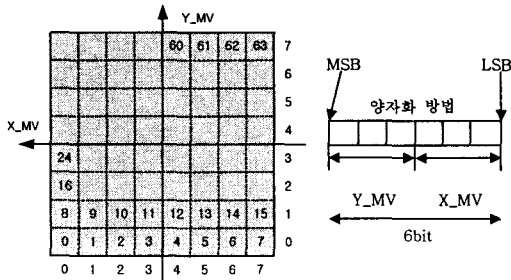


그림6. 64개의 bin으로 이루어진 MV histogram

III. Histogram Memory Interface

두 개(A,B) 의 internal memory(각 288byte)를 이용하여 실시간으로 parsing 되는 정보를 picture 별로 번갈아 가며 bin counting 을 한다. 현재 picture 의 histogram 정보는 memory A 에 read/write 을 하며 counting 이 증가되며 이 때 system 은 이미 저장된 이전 picture 의 histogram 정보가 담긴 memory B 를 read 하여 S/W 에 전하여 준다. 이 때 또 다른 picture 가 시작되면 memory B 를 reset 시킨 후 현재 picture 의 새로운 histogram 정보를 read/write 하며 새로운 counting 을 하고 이 때 시스템은 memory A 의 정보를 read 해 간다.

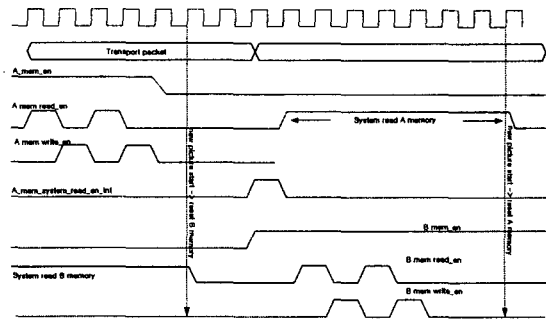


그림7. memory interface timing

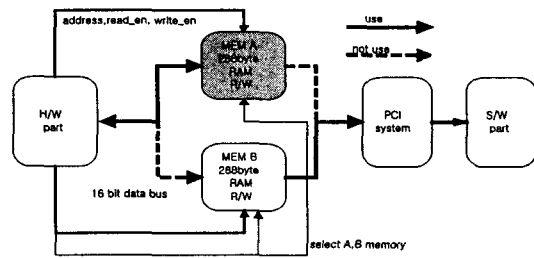
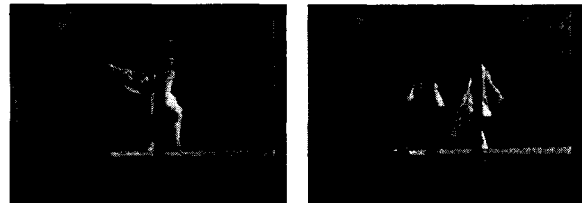


그림8. memory interface block diagram

IV. 모 의 실험



(a)

(b)

(c)

그림9. 장면전환이 발생한 테스트 영상

각각의 image 는 모두 I-picture 로써 DC histogram 을 구하여 이전 picture 와의 변화로 shot change detection 을 수행한다. 테스트 영상 (a)와 (b) 사이에는 약간의 움직임이 있을 뿐이지만 (c)에서는 확실히 shot change 가 일어난 영상이다. 이를 H/W 적으로 만든 DC histogram 으로 비교하면 그림 10 에서 알 수 있듯이 shot change 가 일어났음을 알 수 있다.

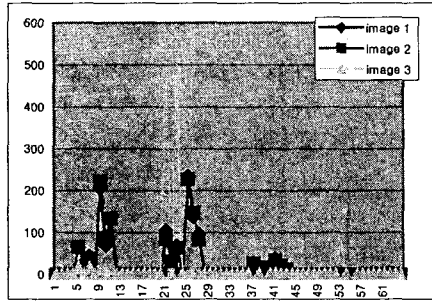


그림 10. image 1,2,3의 DC histogram 비교

아래의 image 는 P-picture 에서의 motion vector 값을 얻어내어 picture 내의 움직임 변화를 알 수 있다. 그림 11 은 두 대의 차의 움직임이 큰 picture 이다.

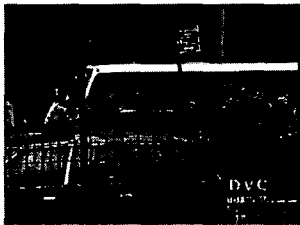


그림 11. P-picture image

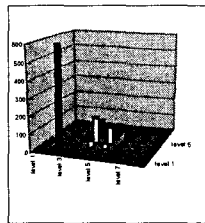
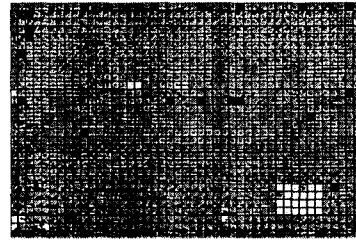


그림 12. MV histogram

그림 13 는 그림 11 을 MB type 으로 다시 나타낸 그림이고 검은 점들의 표시는 motion vector 의 움직임 방향과 그 크기를 나타내는 것이다. 검은 점들의 움직임은 두 대의 차를 중심으로 움직임이 심하게 있었음을 관찰할 수 있다. 이를 근거로 Non Linear Browsing 의 검색속도를 결정하여 일정 배속의 searching 외에 움직임의 변화에 따라 검색속도를 결정하는 algorithm 을 적용하여 보다 많은 상황에 대처할 수 있게 할 수 있다.



Forward MB
 Intra coded MB
 Not coded MB

그림 13. MB_type & motion vector도

V. 결 론

본 논문에서는 DTV 동영상을 H/W 기반의 실시간 검색을 하기 위한 구조에 대해 기술하였다. H/W 구현을 위한 VHDL 코딩에 의한 Simulation 결과와 S/W 적으로 분석한 결과와 일치하였으며 이는 실제 MPEG-2 parsing 부분의 구현에 있어서 많은 부분이 간략화 됨으로 H/W 의 안정성과 속도가 향상되어 실시간 indexing 을 위한 picture 내의 histogram 정보를 가공하는 것이 가능하다는 것을 보여주는 것이다. 이에 마련된 H/W 를 기반으로 이를 ASIC 화 하고 지속적인 upgrade 가능한 S/W 를 구축 하여 DTV 수상기나 Set-Top box 에 적용하여 앞으로 있을 영상시대에 대비해 나가야 할 것이다.

참 고 문 헌

- [1] Ullas Gargi, Rangachar Kasturi, and Susan H. Strayer, "Performance Characterization of Video-Shot-Change Detection Methods", Circuits and Systems for Video technology, vol 10, pp 1-13, Feb 2000
- [2] Bertrand Chupeau, Edouard Francois, "Region-based motion estimation for content-based video coding and indexing", SPIE, vol 4067, pp884-893, June 2000
- [3] 장현성, 김재곤, 김진웅, "비디오 요약을 위한 주요구간 검출 및 표현 방법", 제 12 회 영상처리 및 이해에 관한 워크샵 발표논문집, pp201-205
- [4] M. Berekovic, G. Meyer, Y. Guo, P. Pirsch, "A Multimedia RISC Core for Efficient Bitstream Parsing and VLD", Multimedia Hardware Architectures, Proceedings of SPIE. Vol.3311, pp. 131-141, Jan. 1998
- [5] Sung-Bae Jun, K. Yoon, H.Y. Lee, Dissolve Transition Detection Algorithm Using Spatio-Temporal Distribution of MPEG Macro-Block Types