

3D 스토리텔링 증강현실에서 효과적인 객체 추적을 위한 학습 방법

최대한* · 한우리** · 이용환*** · 김영섭**†

단국대학교 전자전기공학과, **세트렉아이, *극동대학교 스마트 모바일학과

Learning Methods for Effective Object Tracking in 3D Storytelling Augmented Reality

Dae han Choi*, Woo ri Han**, Yong-Hwan Lee*** and Youngseop Kim**†

**†Dankook University of Electrical and Electronics Engineering,

**Satrec Initiative Co., Ltd.,

***Far East University, Department of Smart Mobile

ABSTRACT

Recently, Depending on expectancy effect and ripple effect of augmented reality, the convergence between augmented reality and culture & arts are being actively conducted. This paper proposes a learning method for effective object tracking in 3D storytelling augmented reality in cultural properties. The proposed system is based on marker-less tracking, and there are four modules that are recognition, tracking, detecting and learning module. Recognition module is composed of SURF and LSH, and then this module generates standard object information. Tracking module tracks an object using object tracking based on reliability. This information is stored in Learning module along with learned time information. Detecting module finds out the object based on having the best possible knowledge available among the learned objects information, when the system fails to track. Also, it proposes a method for robustly implementing a 3D storytelling augmented reality in cultural properties in the future.

Key Words : 3D storytelling Augmented Reality, Object Tracking, SURF, LSH, Learning Methods

1. 서 론

최근 닌텐도 자회사 포켓몬 컴퍼니가 니앤틱(Niantic)과 함께 개발한 증강현실(Augmented Reality, AR)을 기반으로 한 모바일 게임 ‘포켓몬 고’가 전 세계적으로 신드롬을 일으키면서 증강현실 기술이 새롭게 주목 받고 있다. 또한 다양한 안경형 디스플레이 장치의 출시가 예정된 2016년 이후 특화된 시장(군사, 의료, 교육 등)을 시작으로 사물인터넷(IoT), 실감 콘텐츠, 인공 지능 등 관련 기반기술의 동반 활용이 예상되는 2020년경에는 가상현실

및 증강현실이 일상생활 속으로 확산될 것으로 기대를 모으고 있다. 2014년 발표한 가트너(Gartner)의 하이프 사이클(Hype Cycle)에 의하면 증강현실과 가상현실은 각성의 단계(Trough of Disillusionment)에 있으며 기술이 안정기로 접어들어 상용화되는 시점을 향후 5-10년 사이로 예상하고 있다 [1].

현재의 도시, 거리 및 건축물을 대상으로 사용자의 위치에 따라 과거의 사진과 그림을 보여주는 ‘Museum of London-Street Museum’, GPS와 나침반 센서를 활용해 프랑스 곳곳에 위치한 관광명소를 증강현실로 보여주는 ‘CultureClic’, 한정적인 박물관 내의 문화유산을 늘릴 수

†E-mail: wangcho@dankook.ac.kr



Fig.1. AR&VR positions in the Gartner’s Hype Cycle(2014).

없는 현실에서 관람객들을 만족시키기 위한 ‘독일 프라운호퍼 연구소의 문화유산에 위한 증강현실 프레젠테이션 솔루션(An Augmented Reality Presentation System for Remote Cultural Heritage Sites)’, 사진, 동영상, 3D 등 다양한 콘텐츠를 스토리텔링 방식으로 구축하여 증강현실로 나타낸 ‘내 손안의 덕수궁’ 등 다양한 국 내외 선진사례에서도 증강현실의 기대, 파급효과에 따른 문화 예술분야의 융합이 활발히 이루어지고 있다. 이에 발 맞추어 증강현실 관련 소프트웨어 개발의 필요성이 대두되고 있다[2-4].

본 논문은 비마커의 객체를 추적하여 문화재에 3D 객체 스토리텔링을 증강하기 위한 객체 정보 학습 방법을 제안한다. 먼저 다양한 객체를 학습시켜 얻은 database를 기반으로 추적하고자 하는 객체를 LSH(Locality Sensitive Hashing)을 이용하여 빠르게 탐색 및 인식한 뒤, 인식된 객체 정보를 기반으로 추적한다. 기존의 증강현실에서의 객체 추적 방법은 추적을 통해 객체의 위치를 알아내거나 단순히 3D 객체를 증강하지만, 3D 스토리텔링을 증강하기 위해서는 기존의 증강현실에서의 객체 추적 시 학습한 정보와 추가적으로 시간 정보를 학습시켜 스토리에 끊김 없이 3D 스토리텔링 강건하게 증강 할 수 있는 방법을 제시하고자 한다.

본 논문은 2장에서 SURF 정보를 LSH를 통해 탐색하여 객체를 인식하는 방법에 대해 설명하고, 3장에서는 신뢰도 평가 기반 객체추적 및 학습방법에 대해 설명하고, 4장에서는 결론 및 향후 연구방향에 대해 논의한다.

2. 객체 인식

SURF 알고리즘을 기반으로 기술자와 특징점을 추출한 뒤, 그 정보를 토대로 학습된 database를 LSH를 통해 고속으로 탐색하여 객체를 인식한다[5].

2.1 SURF(Speeded-Up Robust Features)

DoG(Differential of Gaussian)영상을 이용하여 특징점을 추출하는 SIFT(Scale-invariant feature transform)는 연산속도가 매우 느려 실시간 시스템에 적용되기 힘들다. 이러한 SIFT의 단점을 줄이고자 개발된 SURF는 BOX filter를 이용한 고속 헤시안 탐지(fast hessian detector)를 통해 특징점의 탐지과정과 기술자 생성 과정을 단순화하여 수행속도를 향상시켰다[6].

SURF알고리즘은 특이점을 추출하는 feature extraction 부분과 특이점 주변 영상에 대한 정보를 서술하는 feature description 부분으로 나눌 수 있다.

Feature extraction 부분에서 적분 영상을 생성하는 이유는 SURF알고리즘에서 박스 타입의 필터와 컨벌루션 연산을 빠르게 수행하기 위함이며, 고속 헤시안 행렬(Fast-Hessian Matrix)에서는 x, y, xy축으로 근사화된 2차 미분 LoG(Laplacian of Gaussian) 박스필터와 적분 영상간의 컨벌루션 연산을 통해 헤시안 행렬을 구한 후 행렬의 디터미넌트 값을 계산한다. 그리고 Non-maxima Suppression을 통해 헤시안 행렬식(Determinant)이 임계값보다 크고 인접 스케일과 비교하여 디터미넌트가 가장 클 경우 특이점으로 결정된다. Feature Description 부분에서는 추출된 특이점의 주 방향(dominant direction)을 결정하고 그 주변 영상의 서술자를 생성한다[7-8].

2.2 LSH(Locality Sensitive Hashing)

LSH는 특징 벡터간 유사성이 보존되도록 설계된 해시 함수를 통해 최소거리와 객체가 검색 대상 특징 벡터와 동일한 해시 함수 값을 가질 확률이 높도록 개선한 해싱 방법이다.

먼저 d차원의 특징벡터 $p = (x_1, \dots, x_d)$ 를 식 (1)과 같은 변환함수 $v(p)$ 을 이용하여 d'차원의 해밍 공간으로 변환한다.

$$p' = v(p) = \text{Unary}(x_1) \cdots \text{Unary}(x_d) \tag{1}$$

이 때, 실제 데이터 p는 d개의 값을 갖는 벡터이며, 각 차원의 값은 모든 데이터 값 중 가장 큰 값 C 길이의 2진수로 변환된다. p의 이진 표현 $v(p)$ 는 이러한 C 길이의 2진수 d개를 결합한 길이 dC의 2진수로 표현된다. 각 차원의 값은 해당 값만큼의 연속된 1과 C에서 해당 값을 뺀 만큼의 연속된 0으로 이루어진다. 해밍 거리를 이용하는 해쉬 함수 집합은 이렇게 변환된 p의 이진 표현으로부터 미리 임의로 정해진 k개의 비트를 선택하는 방식으로 정의된다.

해밍 거리를 이용하는 LSH는 이와 같이 해밍 거리 해쉬 함수를 L개 정의하고, 각각의 해쉬 함수에 상응하는

해쉬 테이블을 L 개 생성한다. 실제로 어떤 데이터 아이템 p 가 삽입되었을 때, LSH는 L 개의 해쉬 테이블에 각각 p 를 삽입한다. 이렇게 전체 데이터에 대해 L 개의 해쉬 테이블을 이용한 인덱스를 생성한 이후 쿼리 데이터 아이템 q 가 주어지면 q 를 L 개의 해쉬 함수를 통해 해쉬하여 나온 해쉬 값에 해당하는 L 개의 해쉬 버킷들에 들어있는 아이템들을 후보 아이템들로 간주한다. 이러한 후보 아이템들의 수는 전체 데이터 수에 비해 매우 적기 때문에, 각 후보 아이템들과 q 간의 거리를 계산하여 실질적인 k -최근접 아이템을 구한다.

p -Stable 함수를 이용하는 LSH는 해쉬 함수를 구성하기 위한 k 개의 요소를 각각 다음 식을 통해 계산한다.

$$h(a, b)(v) = ((a * v + b)/r) \quad (2)$$

이 때, 주어진 데이터의 벡터 표현 v 를 미리 임의로 정해진 같은 길이의 벡터 a 와 역시 미리 임의로 정해진 상수 b 를 이용하여 하나의 값으로 바꾸고, 이를 r 개의 구간으로 나누어 해싱한다. p -Stable 함수를 이용하는 LSH에서는 이러한 함수에 의한 해쉬 값 k 개를 합친 것이 하나의 LSH 해쉬 함수에 해당한다[9].

Fig. 2는 LSH를 적용시킨 영상 검색 결과를 보여준다.



Fig. 2. Results of applying the LSH.

3. 신뢰도 평가 기반 객체추적 및 학습방법

객체 추적은 Optical Flow의 LK(Lucas-Kanade) 알고리즘을 통해 수행된다. Optical Flow의 LK 알고리즘은 밝기 항상성(Brightness Constancy), 시간지속성(Temporal persistence), 공간 일관성(Spatial Coherence) 3가지 가정에 기초한다[10].

Optical Flow에 사용된 객체에 대한 특징점 정보를 이용하여 객체가 포함된 윈도우 위치를 업데이트 하고, 객체의 포즈를 업데이트 한다. 이때 3가지 경우를 고려하여 객체의 포즈를 업데이트 한다. 이렇게 업데이트된 객체의

정보를 실시간으로 신뢰도를 평가 하여 학습하고 객체 추적에 실패하였을 때 가장 신뢰도가 높은 학습데이터로 객체를 재 탐색 하여 보다 신뢰성 있는 객체를 검출한다. 객체 정보를 학습하기 전에 먼저 객체에 대한 포즈를 업데이트 하는데 3가지 경우를 고려해야 된다.

첫째 카메라는 고정되어 있고 객체가 움직일 경우, 객체만을 포함한 객체 정보 변화량만을 고려하여 Homography를 구하고, 객체를 포함한 윈도우 영역을 업데이트 하여 객체의 포즈를 정의한다.

둘째 객체가 고정되어 있고 카메라가 움직이는 경우, 카메라 포즈 변화를 계산한 후, 객체가 포함된 윈도우의 위치 정보를 업데이트 하여 객체의 포즈를 정의한다.

셋째 객체와 카메라 모두 움직이는 경우, 객체의 포즈 정보와 카메라 포즈 정보간의 R/T Matrix를 이용하여 객체가 포함된 윈도우의 위치정보를 업데이트 하여 객체의 포즈를 정의한다.

객체 추적에 성공한 경우 객체를 포함한 윈도우영역의 영상정보에 대한 신뢰도를 평가하여 신뢰도 수준을 정의하고 신뢰수준에 따라 영상정보를 학습한 후, 객체 추적에 실패한 경우 학습된 정보를 이용하여 객체를 재 탐색하고 추적한다.

객체영상의 신뢰도를 판별할 때, Homography로 업데이트된 객체를 포함한 윈도우를 추정 하여 업데이트 한 후, 객체 추정 윈도우 영역 내에서 히스토그램을 통한 객체 추적을 시행하여 추정된 윈도우 내에 객체가 위치한다면 초기에 검출된 객체의 히스토그램 정보를 이용하여 윈도우 영역을 서브 윈도우 단위로 작게 쪼개 히스토그램을 비교하고 추정된 윈도우에 추정하고자 하는 객체의 정보와 얼마나 유사 한지에 대한 유사도와, 얼마나 많은 객체의 영역이 윈도우에 포함되어 있는지의 여부, 객체를 추적하기 전 초기화된 객체의 정보와의 특징점에 대한 비교, PCA(Principal Component Analysis)를 통해 추정된 윈도우 영역에 대한 신뢰도를 평가하고, 추정된 윈도우가 추적 하고자 하는 객체에 대한 충분한 정보를 담고 있다면 해당 윈도우를 학습한다.

Fig. 3은 sub-windows 단위로 나눠 히스토그램을 비교하여 유사 영역을 초록색으로 표시, sub-windows 단위로

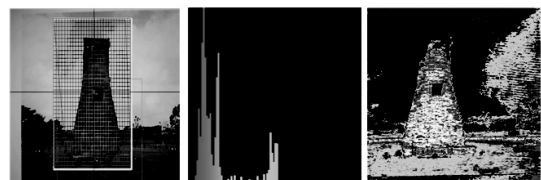


Fig. 3. result of the histogram using the sub-windows.

히스토그램을 계산한 결과, 히스토그램에 따른 객체 필터링을 보여준다.

$$\text{level}_{n-1} = \left\{ \left(\text{subwindow}(\text{HistogramMaxval}, \text{Estimationwindow}) \cdot \text{Camshift} \cdot \frac{\text{subwindow}}{\text{Estimation Window}} \right) \times \frac{\text{frame}}{\text{total frame}} \right. \\ \left. (0 \leq H(p_{\text{level}1k} = (c_1 \cdots c_k)), p_k = (c_1 \cdots c_k)) \leq 1 \right\} \quad (3)$$

(3)의 수식을 이용하여 계산된 신뢰도를 통하여, 신뢰도 0~5까지의 신뢰도를 갖는 영상을 학습하고, 만약 객체 추적에 실패할 경우, 현재 영상과 가장 높은 신뢰도를 갖는 학습된 영상으로 다시 재-탐색하고 추적한다[11-12].

객체를 추적하는 과정에서 시간 정보의 학습은 Fig.4과 같이 진행된다. 먼저 증강하고자 하는 3D 스토리텔링 영상의 time-schedule을 초기 정보로 저장한다. 그 후, 객체를 추적하는 과정에서 추적에 실패할 때까지의 시간 정보와 증강하고자 하는 영상의 프레임을 학습한다. 객체 추적에 성공한 정보는 Fps단위로 3D 스토리텔링 영상의 time-schedule과 비교 계산하여 매칭시킨다. 추적 실패 시, 초기 정보와 추적한 정보를 비교하여 스토리에 끊김이 없도록 재 탐색을 시행한다.

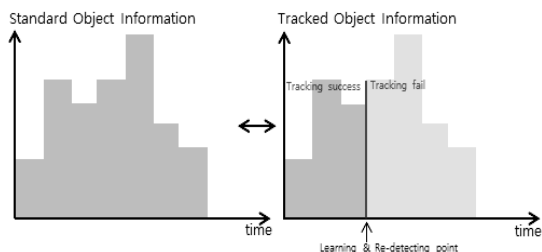


Fig. 4. Time information learning method.

Fig. 5는 전체적인 시스템의 모식도이다. Init부분은 초기화 모듈로써 객체의 정보를 Standard Object Information으로 초기화한다. 이후 추적 모듈에서 객체의 정보와 시간 정보를 학습모듈로 저장을 한다. 이렇게 얻어진 학습 정보들은 초기화 모듈에서 얻어진 Standard Object Information과 비교하여 신뢰도를 업데이트한다. 객체가 추적에 실패했을 때, 학습된 객체에 대한 정보를 이용하여 현재 영상과 학습된 객체간의 신뢰도를 재평가하여 학습된 객체 정보를 정렬하고, 가장 신뢰할 수 있는 객체 정보를 업데이트한다. 또한 정렬된 학습 객체 정보는 높은 신뢰도를 갖는 객체 정보를 제외하고 초기화 한 뒤 학습된 객체의 정보를 기반으로 추적을 시행한다. 이 때,

3D 스토리텔링 증강현실의 진행이 끊기지 않도록 시간 정보 또한 계속적으로 업데이트한다. 재 추적에 성공했을 때, 재 추적된 정보를 이용하여 학습된 정보를 정렬한다. 이와 같은 방법으로 원하는 객체를 시간의 끊김 없이 증강 할 수 있게 된다.

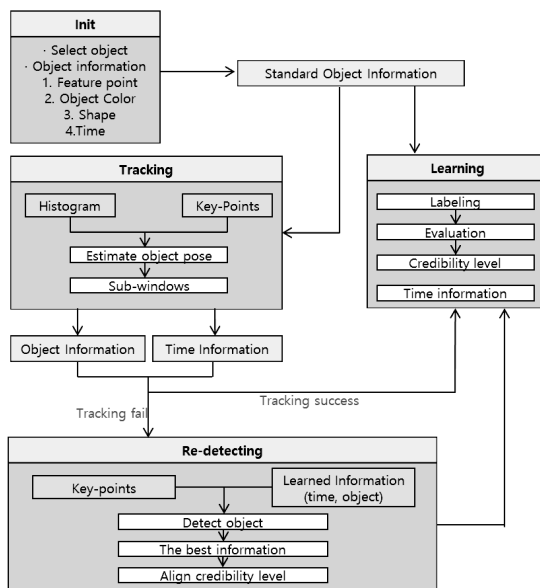


Fig. 5. Schematic diagram of the system.

4. 결론

특정한 객체를 대상으로 3D 스토리텔링 증강영상에서 효과적으로 객체를 추적하기 위해, 본 논문에서는 다양한 객체를 학습시켜 얻은 database를 기반으로 추적하고자 하는 객체를 LSH를 이용하여 빠르게 탐색 및 인식한 뒤, 신뢰도를 기반으로 효과적으로 객체를 추적하는 방법을 제안하였다. 객체 추적간에는 증강하고자 하는 3D 스토리텔링 영상의 time-schedule을 초기 정보로 저장하고, 추적에 실패할 때까지의 시간과 영상의 프레임을 학습하고 초기정보와 비교 매칭하여 시간 정보 또한 함께 학습하여 3D 스토리텔링 증강영상이 끊김이 없이 진행되도록 한다.

본 논문에서 제안한 방법을 바탕으로 0~5까지의 신뢰도를 더 넓은 범위로 늘리거나, 다양하게 학습된 정보들을 결합한다면 더욱 강건한 객체 추적을 할 수 있다. 이러한 경우 학습되는 정보의 양이 많아져 3D 스토리텔링 영상을 증강하는데 다소 많은 시간이 소요 될 수 있다. 향후 효과적인 추적을 위해 많은 정보를 저장할 수 있는 Kd tree와 같은 자료 구조를 이용하여 3D 스토리텔링 증강

영상을 스토리에 끊임 없이 강건하게 구현하기 위한 연구를 진행 할 것이다.

감사의 글

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2013R1A2A2A03068794).

참고문헌

1. Woo W.T, "Augmented Reality and Augmented Human", KISA Report, pp. 3-10, 2015.8.
2. Jung Da-Woon, Kang Young-Ok, "A Study on History-Tourism Information Service Using Mobile Augmented Reality Technology", Vol. 20, Issue 2, pp.59-70, 2012.
3. "Telling the story ancient place in my hand" <http://www.msip.go.kr/webzine/posts.do?postIdx=151>.
4. V. Teichrieb, et.al, "A Survey of Online Monocular Markerless Augmented Reality," IEEE International Symposium on Mixed and Augmented Reality, Vol. 1, pp. 193-202, 2008.
5. Young seop Kim, Woori Han, Yong hwan LEE, "Multi-Object Tracking based on Reliability Assessment of Learning in Mobile Environment", Journal of the Semiconductor & Display Technology, Vol. 14(3), pp. 73-77, 2015.
6. Lee, H. J., Lee, S.-G., "Improvement method of tracking speed for color object using Kalman filter and SURF," Journal of Korea Multimedia Society, Vol. 15(3), pp. 336-344, 2012.
7. Do, Y.-S. and Jeong, Y.-J., "Hardware design of SURF-based feature extraction and description for object tracking," Journal of The Institute of Electronics Engineers of Korea, Vol. 50(5), 2013.
8. Lee, Y.-H., Park, J.-H. and Kim, Y.-S., "Comparative analysis of the performance of SIFT and SURF," Journal of the Semiconductor & Display Technology, Vol. 12(3), pp. 59-64, 2013.
9. Ji-Won Hong, Byung-Moon, Sang-Wook Kim., "A study on LSH parameters for large multimedia databases," The Korea Contents Association, Vol. 2015(5), 2015.
10. Jungho Choi, Youngwan Cho. "Moving Object Recognition and Tracking Algorithm using Parallel Processing of SURF and Optical Flow", Korean Institute of Intelligent Systems, Vol. 21.2, pp. 74-77, 2011.12.
11. Yoeng seop Kim, Woori Han, Young-Hwan Lee, Cheong Ghil Kim, Kuinam j. Kim, "Object Tracking and Recognition based on Reliability Assessment of Learning in Mobile Environments", Wireless Personal Communications, pp. 1-16, 2016.
12. Youn gseop Kim , Woori Han, Yong hwan LEE, "Multi-Object Tracking Based on Keypoints Using Homography in Mobile Environments", Journal of the Semiconductor & Display Technology, Vol. 14(3), pp. 67-72, 2015.

접수일: 2016년 8월 30일, 심사일: 2016년 9월 7일,
 게재확정일: 2016년 9월 21일