

반복적 2D 워핑을 이용한 효율적인 스테레오 비디오 데이터 압축 및 복원

*박일권 **변혜란

연세대학교

[*ikheart@yonsei.ac.kr](mailto:ikheart@yonsei.ac.kr)

Efficient Data Compression and Decompression of Stereo Video by using Iterative 2D Warping

*Park, Il-Kwon **Hye, Ran-Byun

Yonsei University

요약

본 논문은 스테레오 비디오 데이터 압축 및 복원을 위한 그리드(Grid) 기반 2D 워핑 방법을 제안한다. 스테레오 비디오에 대한 데이터 표현 방법으로 깊이지도 및 매쉬(mesh)를 이용한 방법이 주로 사용되어 왔으며 매쉬를 이용한 방법은 두 영상간의 매칭되는 노드를 이용하여 데이터 압축 효율을 높일 수 있다. 그러나, 두 영상에서 매칭되는 노드의 위치를 찾는 것은 매우 어려운 일일 뿐만 아니라 매쉬에 의해서 워핑된 영상과 목적이 되는 스테레오 영상의 좌측 또는 우측 영상간의 왜곡이 불가피하다. 따라서 이러한 왜곡을 보정하기 위하여 잔여영상(Residual image) 정보를 추가로 요구하게 된다. 제안된 논문은 이러한 잔여영상 정보를 최소화 하기 위하여 반복적으로 2D워핑을 수행하며 최적화된 워핑 영상을 생성함으로써 목적영상과의 오차를 최소로 유지하여 추가정보인 잔여 영상의 데이터 용량을 최소화 한다. 전체영상에 대하여 2D워핑을 수행하며 각각의 노드를 변경하는 것은 많은 비용을 감수해야 하기 때문에 오차영역에 대하여 지역단위로 분할하고 단계적으로 최적화를 이루는 분할정복 방법을 사용하였다. 본 논문의 실험에서는 스테레오 영상에 대하여 각각의 신호대 잡음비(PSNR)를 통해 제안한 방법의 품질을 평가하였을 뿐만 아니라 기존의 메쉬 기반한 방법과 깊이지도를 이용한 방법과의 데이터량을 비교하였다. 실험 결과를 통하여 제안한 방법의 데이터 압축의 효율성 및 품질의 우수성을 확인하였다.

1. 서론

최근 소비자의 시각적 욕구와 기술의 발전으로 3차원 입체 영상에 대한 관심 및 수요가 증가하고 있다[1]. 특히, 이와같은 3차원 입체 콘텐츠는 과학적 연구, 교육, 오락, 광고를 목적으로 제작되고 있을뿐만 아니라 3차원 방송을 위한 비디오 콘텐츠 형식으로 자리잡고 있다 [7]

.하지만 스테레오 비디오 콘텐츠는 기본적으로 좌측과 우측의 한쌍의 영상으로 구성되기 때문에 스테레오 비디오 콘텐츠를 위한 저장공간 및 전송량에 대한 물리적 한계를 가지고 있다. 기존의 연구에서는 주로 데이터 압축을 위해 좌측 영상과 우측영상 복원을 위한 그레이스케일(Grayscale) 시차맵을 사용하였다 [3]. 또 다른 연구는 왼쪽 영상과 산재된 시차맵(Sparse disparitymap)에 대한 메쉬 모델링을 통한 오른쪽영상 복원방법을 사용하였다 [1, 5]. 반면에, 본 논문에서는 왼쪽영상에 대하여 제어점을 변경해가며 반복적인 2차원 워핑기술을 적용하여 오른쪽영상을 복원하는 그리드 기반(Grid-based)의 접근방법을 제안한다.

시차정보에 의한 스테레오 비디오 표현은 왼쪽과 오른쪽 두영상을 이용하는 대신 한쪽 영상과 그레이스케일의 시차맵을 이용한다. 시차맵은 왼쪽 영상과 오른쪽 영상간의 시차정보를 최대 255값의 그레이스케일 형태로 표현된다. 데이터 크기는 원본 컬러 영상의 30%정도를 차지한다. 또한 시차맵은 기존의 영상 압축 알고리즘을 사용하여 압축할 수 있는 장점이 있다. 그러나 이러한 접근법은 여전히 많은 데이터용량을 요구할뿐만 아니라 스테레오 영상에서 일치하는 픽셀관계를 추정하기 어렵기 때문에 정교한 시차맵 생성이 어렵다. 특히 가려짐 영역, 반복되

는 텍스처 영역 그리고 동일한 텍스처값으로 이루어진 영역등에 대한 추정은 많은 왜곡된 정보를 생성한다 [8]. 이런 문제점을 극복하며 고품질의 영상 복원을 위해 기존의 접근법에서는 잔여 텍스처 정보(Residual texture image)를 추가적으로 제공하였다.

한편, 메쉬 기반의 접근 방법은 깊이지도 기반의 접근 방법보다 좀더 축약된 스테레오 비디오 압축을 제공한다. 왜냐하면 메쉬모델링에 의한 영상은 반대쪽 영상 생성 복원을 위한 단지 소수의 메쉬 제어점만 저장하면 된다[1]. 그러나 메쉬기반의 접근방법 또한 가려짐 영역에서는 일치 않는 영상 왜곡을 야기 시킨다. 이러한 문제를 극복하기 위해 몇몇 방법에서는 왜곡된 영역에 에러값이 최소가 되도록 제어점을 추가하는 방식으로 메쉬 모델링을 수행하였다. 시차맵 기반의 접근법과 마찬가지로 메쉬 기반접근법 역시 고품질의 영상 복원을 위해서는 잔여 텍스처 정보(Residual texture image)를 제공해야만 한다.

본 논문에서 왼쪽영상에 대하여 반복적인 2차원 워핑 기술을 적용하여 오른쪽영상을 복원하는 방법을 제안한다. 2차원 워핑을 위해서는 소수의 제어 점들만을 요구하기 때문에 메쉬 기반의 접근법과 마찬가지로 데이터 크기를 줄일 수 있다. 기존의 메쉬접근법과 제안한 방법의 주요 차이점은 델로네(Delaunay) 알고리즘과 같은 별도의 메쉬 생성 알고리즘을 사용하지 않을 뿐만 아니라 기존의 방법은 산재된 시차맵을 이용하는 반면 제안한 방법은 시차맵을 전혀 사용하지 않는다.

본 논문의 기여는 반복인 2차원 워핑방법을 이용하여 최적화된 워핑결과를 생성함으로써 최소화된 잔여 텍스처 영상을 생성하도록 제안

하였다. 복원된 영상안에 나타나는 인공적인 요소를 최소화하기 위해서 에러맵(Errormap)을 작성하고 에러맵에 속한 영역에 대해서만 제어점의 위치를 변경해 가며 반복적인 2D워핑을 수행하고 그 결과를 목적영상과 비교하여 오차가 최소가 되는 제어점 위치를 추출한다..

본 논문의 구성은 2장에서 제안하는 알고리즘에 대한 전체 흐름도를 나타내며 3장에서는 제안한 반복적 2D워핑 방법을 위한 특징점 추출 및 제어점 설정에 관한 방법을 설명하였으며 3장에서는 반복적 2차원 워핑 알고리즘에 대해서 설명하고 4장에서는 영상의 품질 및 데이터량에 따른 실험 결과를 나타내며 마지막으로 본 논문의 결론 및 향후 연구에 대해서 정리하였다.

2. 제안하는 스테레오 비디오 압축 알고리즘

제안하는 스테레오 비디오 압축 알고리즘은 그림 1과 같이 크게 전역 처리 부분과 블록단위의 지역처리로 나눌 수 있으며 기능상으로는 제어점 등록 모듈과 반복적 2D워핑 모듈로 구성되어 있다. 영상 워핑을 수행하기 위해서는 제어점 추출 및 등록이 중요하며 좌측 영상과 우측영상에 매칭되는 제어점을 찾기 위한 방법을 제안하였다. 한편 등록된 제어점을 기준으로 그리드를 형성하고 영상워핑을 수행하여 반대편 영상을 복원한다. 이때 목적 영상과 복원된 영상간의 에러를 구해 에러맵을 구성하고 에러맵에 따라 영상을 블록단위로 나누고 각각의 블록내의 제어점을 수평으로 이동해 가며 블록단위로 영상 워핑을 수행한다. 에러값이 최소가 되는 제어점 위치를 등록하고 차례로 최적화되지 않은 나머지 블록에 대하여 위와 같이 최적화된 제어점의 위치를 추출하게 된다. 이와 같이 분할 정복방법을 이용하여 에러값이 높은 순서대로 각각의 블록에 대해 반복적으로 최적화된 워핑 결과를 얻어내기 때문에 최종적으로 전체 영상에 대한 최적화된 워핑 결과를 얻을 수 있다.

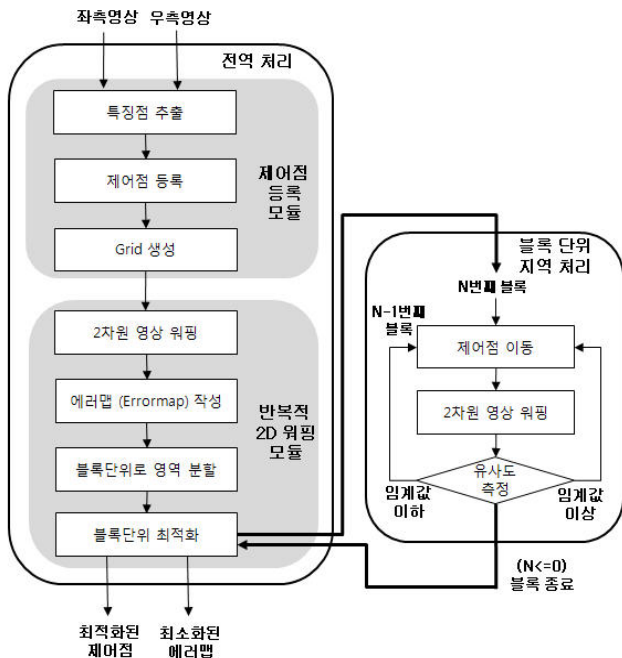


그림 1. 반복적 2D 워핑을 이용한 스테레오 비디오 압축 알고리즘

3. 특징 추출 및 제어점 설정

일반적으로 2차원 워핑을 수행하기 위해서는 원본 영상과 목적영상

에 대한 제어점 설정의 정확성에 따라 결과영상의 품질이 의존적이다. 본 논문에서는 그림 2와 같이 자동적인 제어점 추출 및 설정을 수행하기 위하여 원본 영상과 목적영상에 Harris corner 검출 알고리즘을 이용하여 특징점을 추출하고 Canny edge 추출 알고리즘을 이용하여 객체의 경계영역 정보를 융합하여 객체의 경계영역을 따라 제어점들이 생성하도록 하였다 [4].

2차원 워핑에 의한 영상 왜곡을 줄이기 위해서는 제어점들이 객체의 경계영역 및 강한 에지라인을 따라 생성되어야만 한다. 워핑에 의한 영상 왜곡은 제어점들 사이의 텍스처 정보를 보간하고 리샘플링하기 때문에 발생함으로 경계영역이나 강한 에지라인이 제어점들 사이에 존재하면 여러개의 에지라인이 리샘플링 과정에서 생성되어 눈에 띄게 왜곡되어 나타나기 때문이다.

Harris corner 검출 알고리즘에 의해 추출된 특징점은 영상내에 에지가 교차하는 모퉁이영역을 주로 추출하며 그러한 모퉁이영역은 조명 변화나 블러링된 영상에서도 비교적 강인하게 추출되기 때문에 특히 추적알고리즘을 위한 좋은 특징점 추출 알고리즘으로 알려져 있다 [4]. 하지만 명확한 객체의 경계영역임에도 불구하고 모퉁이 영역이 아니기 때문에 경계영역에 대한 특징점 추출이 이루어지지 않기 때문에 워핑을 위한 제어점 추출로 사용하기 위해서는 보완되어야 할 필요가 있다. 따라서, 본 논문에서는 Canny edge 추출 알고리즘을 이용하여 영상 전체에 edge성분을 검출하고 실험적인 임계값을 통해 객체의 경계영역 및 강한 에지라인을 추출하였다. 이렇게 추출된 에지정보는 수평 방향과 수직 방향으로 20픽셀 단위로 하나의 특징점과 같이 표시하도록 하였다. 이러한 edge정보와 모퉁이 정보를 융합하여 전체적인 특징점 정보로 이용하였다. 또한 전체 영상에 대한 특징점 추출인 경우 특정 영역에 특징점들이 몰리는 현상이 나타나기 때문에 블록단위로 나누고 각각에 대하여 특징점 추출을 수행함으로써 영상 전체에 골고루 특징점들이 분포하도록 하였다.

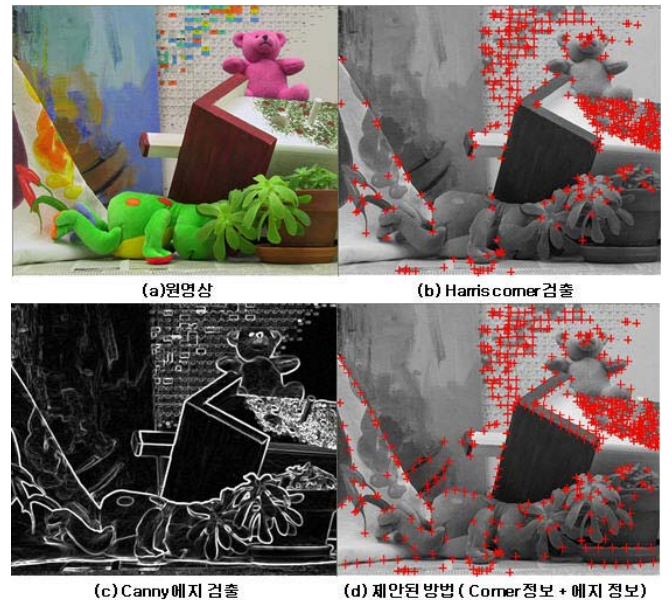


그림 2. Teddy영상에 대한 특징점 추출 결과

추출된 특징점들은 영상 워핑을 위한 제어점들이 되기 위해서 각각의 특징점을 중심으로 블록 매칭 방법을 통하여 특정 임계값 이상의 특징점만을 제어점으로 등록시키고 나머지 특징점들은 버리게 된다.

4. 반복적 2차원 워핑 알고리즘

자동으로 등록된 제어점을 이용하여 좌측영상에 대하여 2차원 영상 워핑을 수행함으로써 복원된 오른쪽 영상과 목적영상인 오른쪽 영상간에 화소단위로 특정 임계값 이상의 오차가 있는 픽셀 정보만을 담고 있는 에러맵을 생성하게 된다. 이렇게 생성된 에러맵은 각각의 오차화소에 대한 원 영상의 RGB 컬러값을 기록하고 나머지 일치하는 화소들은 검은 영역으로 표시하게 된다. 그림 3은 제안한 반복적 워핑 알고리즘을 통하여 에러맵의 크기를 최소화 함으로써 최적화된 영상복원 결과를 나타낸다.

그림3의 초기 에러맵과 같이 초기 워핑결과는 이러한 오차화소가 많기 때문에 에러맵의 용량이 상당히 크게 나타난다. 에러맵의 용량이 크면 클수록 스테레오 비디오를 복원하는 과정에서 추가적으로 영상 품질을 보정하기 위해 사용되는 데이터가 증가하기 때문에 영상압축의 의미를 상실하게 된다.

따라서 본 논문에서 제안한 반복적 2차원 워핑 알고리즘은 에러맵을 참조하여 각각의 블록별로 최적화된 워핑을 수행하여 에러맵의 용량이 최소가 되도록 한다. 에러맵을 참조로 각각의 블록별로 에러가 발생한 블록에 대하여 제어점을 수평방향으로 이동시키며 영상 워핑을 수행하고 복원된 블록과 목적영상의 원 블록에 대하여 에러값을 측정하여 에러값이 최소가 되는 위치의 제어점을 등록한다. 전체 영상에 대하여 반복적으로 제어점을 수정해가며 워핑을 수행하는 것은 현실적으로 어렵기 때문에 분할 정복 방법을 기반으로 에러맵 정보를 참조하여 에러량이 특정 임계값 이상인 영역에 대해서만 반복적으로 제어점을 수정하고 워핑 결과를 비교해가며 최적화된 워핑을 유도한다.

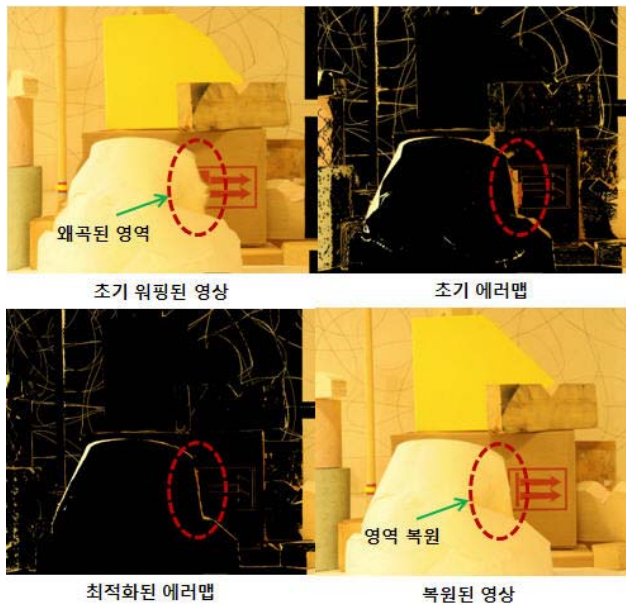


그림 3. 반복적 워핑을 통한 최적화된 영상 복원

5. 실험 및 분석

제안한 알고리즘의 성능평가를 위하여 [2]에서 제공하는 스테레오 영상에 대하여 PSNR을 이용한 영상 품질과 제어점 수에 따른 데이터량을 측정하였다.

그림 4는 Teddy 영상과 Sawtooth 스테레오 영상에 대한 기존의 메쉬 기반 깊이도 표현 알고리즘들과 제안한 알고리즘의 영상품질을 제

어점 수와 PSNR 측면에서 비교하였다. PSNR값은 수식 1에 의해서 획득되어지며 원본 영상과 동일한 경우는 50 dB값을 나타내며 일반적으로 30dB이상의 영상인 경우는 원본 영상과 거의 유사할 정도로 높은 품질을 나타낸다.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) [dB] \quad (1)$$

$$MSE = \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N [f_{original}(x, y) - f_{in-between}(x, y)]^2 \quad (2)$$

그림4의 실험 결과로부터 다른 메쉬기반 알고리즘들은 제어점 수가 600개 이하일 경우 현저하게 낮은 PSNR값을 나타내는 반면 제안한 알고리즘은 상당히 높은 값을 유지하였다. 또한 제어점 수가 증가할수록 PSNR값도 점차 증가하였다. 제안한 알고리즘은 제어점 수가 증가할수록 PSNR값도 같이 증가하였을 뿐만 아니라 전반적으로 모든 제어점 수에 대하여 좋은 성능을 보였다. 특히, [5]에서 제안한 노드 추가방법보다 근소한 차이로 좋은 성능을 보였다. 제안한 방법은 제어점 수의 증가 없이 제어점의 위치를 바꿔가며 PSNR값이 가장 높게 나오는 제어점의 위치를 찾기 때문에 실험 결과와 같이 제어점 수가 같은 경우 기존의 제어점 수에 의존하는 알고리즘보다 높은 PSNR값을 보장한다.

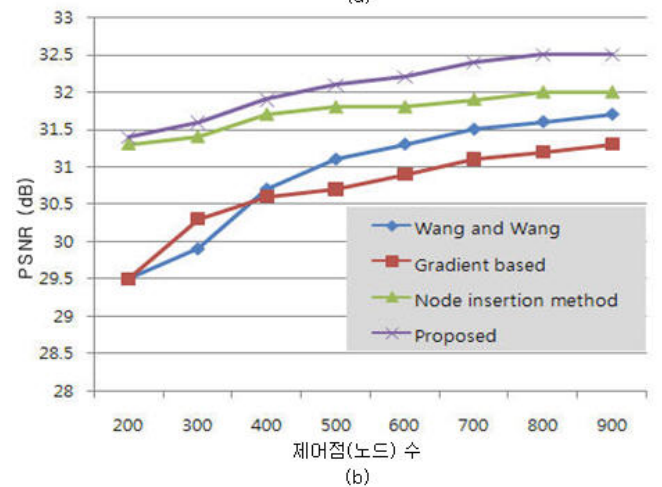
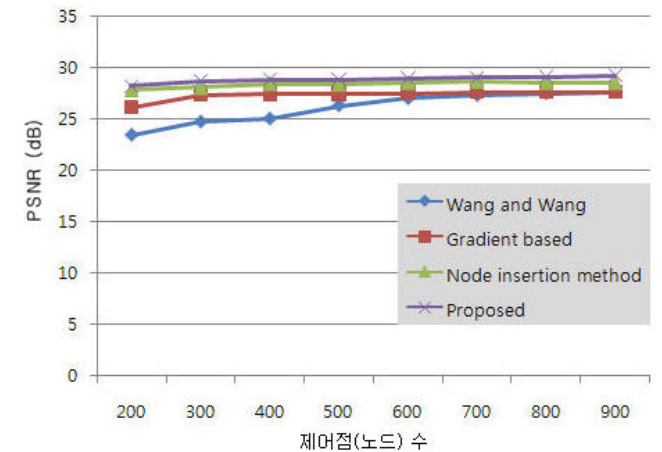


그림 4. 기존의 메쉬기반 알고리즘들과 제안한 방법의 노드 수에 따른 영상 품질 비교 (a) Teddy 영상 (b) Sawtooth 영상

따라서 제안한 알고리즘은 적은 제어점을 가지고 높은 PSNR값을 갖는 고품질의 영상을 복원할 수 있기 때문에 데이터량을 효율적으로 줄일 수 있을 뿐만 아니라 PSNR값이 높은 복원영상은 결국 복원 영상을 보정하기 위한 에러맵의 화소량이 줄어들기 때문에 제어점 수와 에러맵을 동시에 줄일 수 있는 특징을 가지고 있다.

6. 결론 및 향후 연구

본 논문은 스테레오 비디오 데이터 압축 및 복원을 위한 효율적인 그리드 기반 2D 워핑 방법을 제안하였다. 제안된 알고리즘은 영상복원을 위한 잔여영상 정보를 최소화하기 위하여 영역별로 제어점을 변경해 가며 최적화된 2차원 워핑결과를 유도한다. 전체영상에 대하여 2차원 워핑을 수행하며 각각의 노드를 변경하는 것은 많은 비용을 감수해야 하기 때문에 오차영역에 대하여 지역단위로 분할하고 단계적으로 최적화를 이루는 분할정복 방법을 사용하였다. 본 논문의 실험에서는 CG영상과 실사로 촬영된 스테레오 영상에 대하여 각각의 신호대 잡음비를 통해 제안한 방법의 품질을 평가하였을 뿐만 아니라 기존의 메쉬 기반한 방법과 깊이지도를 이용한 방법과의 데이터량을 비교하였다. 실험 결과를 통하여 제안한 방법의 데이터 압축의 효율성 및 품질의 우수성을 확인하였다.

감사의 글

본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 육성·지원사업 (IITA-2008-C1090-0801-0017)의 연구결과로 수행되었음.

참고문헌

- [1] H. Fan and K. N. Ngan, "Disparity map coding based on adaptive triangular surface modeling," *Signal Processing: Image Communication*, vol.14, no.2, pp.119 - 130, 1998.
- [2] "<http://vision.middlebury.edu/stereo/>"
- [3] R. Krishnamurthy et al., "Compression and transmission of depth maps for image based rendering," in *Proc. Int. Conf. Image Processing*, Oct. pp.828 - 831, 2001.
- [4] F. Mohanna and F. Mokhtarian. "Performance evaluation of corner detection algorithms under similarity and affine transforms", *BMVC(2005)* pp.353 - 362. 2005.
- [5] J. H. Park and H. W. Park, "A mesh-based disparity representation method for view interpolation and stereo image compression", *IEEE Trans. on Image processing*, vol.25, pp.1751-1762. 2006.
- [6] J. H. Park and H. W. Park, "Fast view interpolation of stereo images using image gradient and disparity triangulation," *Signal Processing Image Communication*, vol.18, pp.401 - 416, 2003.

- [7] I. Park, M. Kim, H. K. Kim and H. Byun, "Interactive multi-view video adaptation for 3DTV", *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, pp.89-92, 2008.
- [8] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. Journal of Computer Vision*, vol.47, pp.7 - 42, 2002.