

영상검색을 통한 초해상도 기법

한유덕*, 이준영**, 권인소**

*한국과학기술원 미래자동차학제전공

**한국과학기술원 전기 및 전자공학과

e-mail : {ydhan, jylee}@rcv.kaist.ac.kr

Super-Resolution using Image retrieval

Yudeog Han*, Joon-Young Lee**, In So Kweon**

*Dept. of Future Vehicle, KAIST

**Dept. of Electrical Engineering, KAIST

요 약

본 논문에서는 많은 양의 외부 데이터를 활용할 수 있는 예제기반 초해상도(example-based super-resolution) 방법을 보다 효율적으로 할 수 있는 예제선정과 그를 위한 최적화기반의 방법론을 제안한다. 외부 데이터베이스 전체에 의존하는 것이 아니라, 예제선정을 위해 영상검색(image retrieval) 방법을 도입하여 입력 영상과 관련 있는 영상을 외부 데이터베이스로부터 찾고 영상들로부터 초해상도 영상을 얻는다. 기존의 방법은 외부 데이터베이스를 모두 사용하기 때문에 입력영상에 불필요한 정보들이 복원되어 초해상도 결과의 질을 저하시킨다. 하지만 제안하는 방법에서는 영상검색을 통해 불필요한 정보들을 미리 제거하여 좋은 결과를 얻을 수 있다. 또한 외부 데이터베이스를 크기에 상관없이 검색된 몇 장의 영상을 사용하기 때문에 기존의 방법에 비해서 속도가 향상되었다.

1. 서론

디지털 카메라, 센서의 발전에도 불구하고 여전히 해상도 향상에 대한 요구가 존재한다. 따라서 저해상도 영상으로부터 고해상도 영상을 얻는 초해상도 방법이 꾸준히 연구되고 있으며, 다양한 분야에서 사용되고 있다. 이러한 초해상도 방법은 크게 2 가지로 분류할 수 있다. 하나는 다중 영상 초해상도(multi-image super-resolution) 방법이고, 다른 하나는 예제기반 초해상도 방법[2]이다.

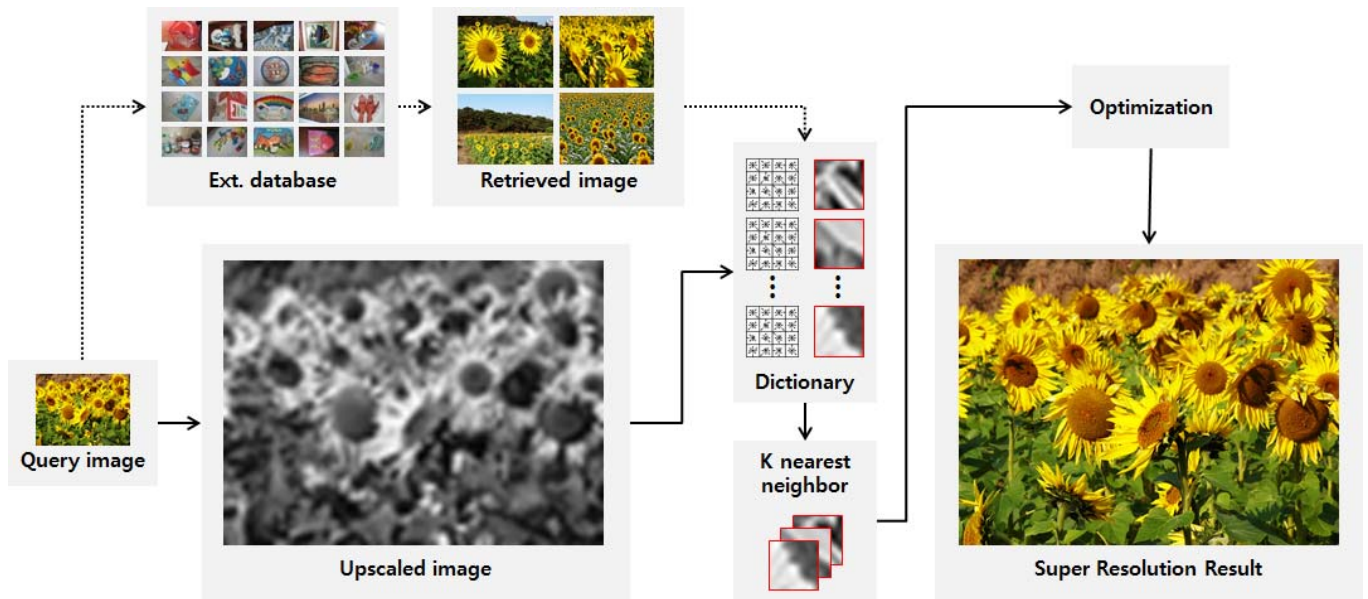
예제기반 초해상도 방법[2]에서는 고해상도 정보를 얻기 위해 외부 데이터베이스를 사용하였다. 입력영상으로부터 추출한 저해상도 패치와 유사한 저해상도 패치들을 외부 데이터베이스로부터 찾는다. 이 저해상도 패치들에 대응되는 외부 데이터베이스의 고해상도 패치들을 이용하여 초해상도를 하는 방법이다. 이 예제기반 초해상도 방법에서는 외부 데이터베이스로부터 고해상도 영상 정보를 얻기 때문에 외부 데이터베이스의 크기가 이 알고리즘의 성능을 좌우하는 가장 핵심이 된다. 하지만 외부 데이터베이스의 크기가 증가함에 따라 속도는 현저하게 느려지게 된다.

Glasner et al.[3]은 다중 영상 초해상도 방법과 예제기반 초해상도 방법을 통합하였다. 이 방법에서는 자연 영상 내에 패치가 중복적으로 나타난다는 자기 유사성을 이용하였다. 외부 데이터베이스를 사용하지 않고, 입력영상 내에 다양한 크기로 존재하는 유사한 패치들을 모은다. 이 다양한 크기의 패치들을 다중 영상 초해상도 방법에서 사용되는 서브픽셀 정렬을

이용해서 한 장의 영상만을 이용하여 초해상도 방법을 적용하였다. 예제기반 초해상도 방법에서 비슷한 성능을 내기 위해서는 상당히 많은 양의 영상들이 외부 데이터베이스로 필요로 하게 된다. 하지만 Glasner et al.[3]은 외부 데이터 베이스를 사용하지 않고 입력 영상만을 사용하기 때문에 속도가 훨씬 빠르다.

Sun et al.[7]은 텍스처 정보를 잘 복원하기 위해서 외부 데이터베이스를 영상이 아닌 영상에서의 텍스처 세그먼트들로 구성하였다. 이 텍스처 세그먼트들을 이용하여 입력영상의 고주파 영역의 디테일을 향상시켰다. 또한 예제기반 초해상도 방법[2]에서 불연속적인 마르코프 모델이 아니라 연속적인 에너지 함수를 최적화하는 방법을 이용함으로써 초해상도 결과 영상의 아티팩트(artifacts)를 최소화 하였다.

본 논문에서는 웹에서 존재하는 다양한 영상들로부터 원하는 영상을 찾기 위한 영상검색 알고리즘을 최대한으로 활용한 방법을 제안한다. 입력영상 외에 다른 다양한 영상들을 활용하기 위해서 외부 데이터베이스를 사용하지 않고 자기 유사성을 사용하는 Glasner et al.[3]의 접근방법과 다르게 외부 데이터베이스를 그대로 유지하였다. 외부 데이터베이스로부터 영상검색을 통해 입력영상과 유사한 몇 개의 영상을 찾고, 이를 사용함으로써 병목 현상의 주된 원인이 되는 외부데이터베이스의 크기에 대한 문제를 해결하였다. 영상검색을 통한 한번의 필터링 과정만으로도 기존의 예제기반 초해상도 방법의 성능이 개선되는 것을 실험으로 검증하였다.



(그림 1) 제안하는 예제기반 초고해상도 방법.

2. 영상검색

영상검색에는 기본적으로 영상과 영상 사이의 정합관계 혹은 영상 속 물체 사이의 정합관계를 정의하기 위해 특징점 기반의 영상 표현방법을 많이 사용되어 왔다. 그러나 특징점을 이용해 추출한 모든 특징점 정보를 사용하게 되는 경우, 데이터량이 너무 많아 자원과 시간이 많이 소요된다. 키워드 기반 텍스트 분류기법의 아이디어를 모방한 Csurka et al. [7]의 코드북 기반 영상 분류방법이 많이 사용되고 있다. 기본적인 프레임워크는 다음과 같다[7][8].

학습 단계에서는 학습영상들에 대한 특징점을 추출하고 기술한다. 이렇게 뽑은 영상기술자를 이용하여 어휘트리(vocabulary tree)[9]를 통해 시각단어(visual words)를 생성한다. 이 시각 단어를 모아서 코드북(codebook)을 만든다. 코드북을 이용하여 각각의 학습 영상들에 대한 Bag-of-Words 히스토그램을 생성한다.

검색 단계에서 입력영상에 대해서 특징점을 추출하고 코드북을 이용하여 똑같이 Bag-of-Words 히스토그램을 생성한다. 학습 영상들에 대한 Bag-of-Words 히스토그램과 정합하여 가장 가까운 영상들을 찾는다. 학습 단계에서는 학습영상들에 대한 코드북을 만드는 과정에서 많은 시간이 소요되지만, 검색 단계에서는 영상을 학습단계에서 만들어놓은 Bag-of-Words 히스토그램을 사용하여 매우 빠르게 찾아낼 수 있다.

3. 제안하는 예제기반 초해상도

제안하는 초해상도 방법은 입력 영상이 주어졌을 때, 입력 영상과 유사한 영상들을 외부 데이터베이스로부터 찾는다. 이 유사한 영상들로부터 패치 사전을 만들고, 입력 영상의 고주파 영역을 복원할 패치들을 찾는다. 최적화 방법을 통해 최종적으로 초해상도로 복원된 영상을 얻는다. (그림 1)

영상검색을 통해 얻은 영상으로부터 패치 사전을 만드는 단계를 학습 단계(training step)이라고 한다. 학습

단계에서는 기존의 예제 기반 초해상도 방법과 같이 외부 데이터 베이스에 있는 고해상도 영상들의 영상 크기를 줄인다. 이 영상을 다시 원래 크기로 복원하면 원래 고해상도 영상의 고주파 성분이 제거되어 흐린(blur) 영상이 된다. 고해상도의 영상과 흐린 영상을 가지고 패치 사전(dictionary)을 다음과 같이 만든다. 고해상도 영상에서 일정한 간격으로 작은 크기의 패치를 추출한다. 흐린 영상에서는 기존의 방법과 다르게 일정한 간격으로 Dense-SIFT[7] 영상기술자를 추출하여 고해상도 패치와 함께 패치 사전을 만든다. 흐린 영상에서 패치가 아니라 Dense-SIFT 를 사용하는 이유는 영상검색을 통해 얻은 유사한 영상 내에 부분적으로 변형되어있는 정보들을 최대한 활용하기 위함이다.

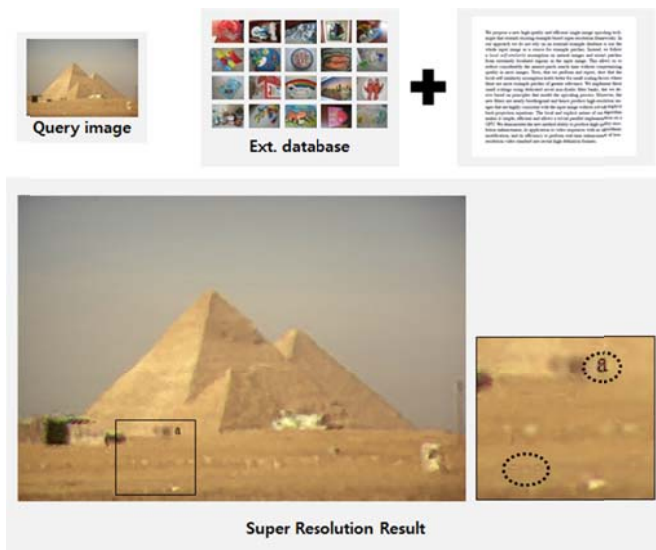
학습 단계에서 만든 패치 사전에서 입력영상의 고주파 성분을 복원할 패치를 찾고, 이로부터 초해상도 결과를 얻는 단계를 합성 단계라고 한다. 합성 단계에서는 입력 영상이 들어오면 Bicubic 알고리즘을 통해 입력 영상의 크기를 키운다. 이렇게 보간된 영상에서 일정한 간격으로 Dense-SIFT 영상기술자를 추출하여 사전에서 kd 트리를 이용하여 k-최근접 이웃을 찾는다. 이렇게 k-최근접 이웃을 찾는 이유는 저해상도 영상기술자 정합을 통해 찾은 고해상도 영상 패치를 가지고 정확하게 입력 영상을 복원할 수 없기 때문이다. 이렇게 영상기술자 정합을 통해 얻은 고해상도 패치들은 복원할 입력 영상의 부분의 방향과 맞지 않게 된다. 입력영상의 특징점의 주된 방향과 맞춰주기 위해서 회전을 시켜주어야 한다. 이 고해상도 패치들을 하나 하나 회전하면 연산량이 상당히 많기 때문에 속도가 상당히 느려지게 된다. 이를 방지하기 위해서 미리 테이블을 만들어 놓고, 테이블을 이용하여 빠르게 회전시킬 수 있다.

마지막으로 회전된 고해상도 패치들을 이용하여 최종 초해상도 영상을 얻기 위해서 최적화를 한다. 최

적화를 하는 방법은 기존의 예제 기반 초해상도방법에서는 마르코프 모델을 사용하는 방법[2]과 Sun et al.[7]에서 제안한 연속적인 에너지 함수를 이용하는 방법으로 나누어진다. 제안한 방법에서는 고해상도 패치들을 회전시켰기 때문에 양자화 에러(quantization error)가 생기게 된다. 이러한 양자화 에러 포함한 채로 마르코프 모델을 사용하여 최적화를 하면 결과 영상에 아티팩트들이 생기는 것을 실험을 통해 보았다. 이러한 문제는 연속적인 에너지 함수를 최적화하는 방법으로 해결할 수 있다.

4. 실험 결과

컬러 영상에 대해서 초해상도 방법을 수행하기 위해서는 RGB 채널로 구성된 컬러 영상을 YCbCr 채널로 변환한다. 이렇게 YCbCr 채널로 변환하는 이유는 영상의 Cb, Cr 채널은 저주파 성분을 주로 가지고 있기 때문이다. Cb, Cr 채널은 Bicubic 알고리즘을 이용하여 보간하며, Y 채널에만 초해상도 방법을 적용하면 된다. Y, Cb, Cr 각 채널에서 얻어진 결과를 합쳐서 RGB 채널로 재 변환하면 컬러 영상에 대한 초해상도 복원된 결과를 얻을 수 있다.



(그림 2) 입력영상에 불필요한 영상이 포함된 경우.

컬러영상에 대해서 2 가지 실험을 하였다. 첫 번째는 외부 데이터베이스에 포함된 입력영상에 불필요한 영상이 결과에 미치는 영향을 알아보기 위한 실험을 하였다. (그림 2)와 같이 외부 데이터베이스에 텍스트 영상을 포함하였다. 예제기반 초해상도 방법을 수행한 결과 복원된 영상의 에지 부근에 텍스트가 있는 것을 볼 수 있다. 이렇게 외부 데이터에 입력영상에 관련이 없는 영상이 초해상도 결과에 안 좋은 영향을 미칠 수 있다는 것을 실험을 통해 검증하였다. 하지만 영상검색을 사용하는 경우 이러한 영상들을 사전에 걸러낼 수 있게 된다.

두 번째는 영상검색과 회전 정보의 사용 유무, 합성단계에서의 최적화 방법에 따른 실험을 하였다. (그림 3, 4)에서 보듯이, 영상검색의 사용만으로도 결과가

향상되는 것을 볼 수 있다. 회전 정보를 사용했을 경우 사용하지 않았을 경우보다 고주파 영역의 디테일이 잘 복원되지만, 아티팩트들 또한 생겨나는 것을 볼 수 있다. 이러한 아티팩트는 불연속적인 마르코프 모델이 아니라 연속적인 에너지 함수로 모델링을 하여 최적화를 함으로써 해결할 수 있다.

기존 예제기반 초해상도 방법에 영상검색을 사용하여 외부 데이터베이스로부터 입력영상과 비슷한 영상들을 찾고, 그 영상들을 이용하여 초해상도를 적용한 결과 보다 적은 학습영상으로 보다 좋은 결과를 얻었다.

5. 결론

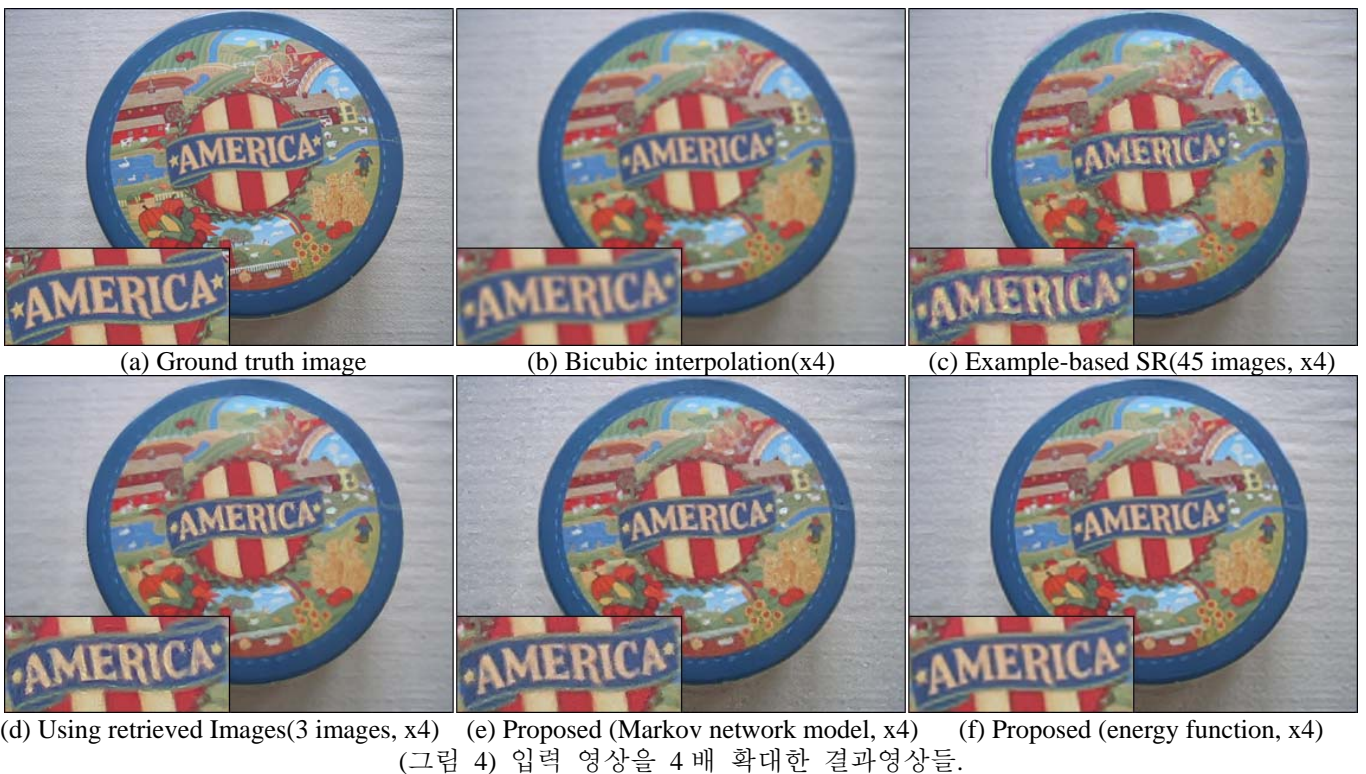
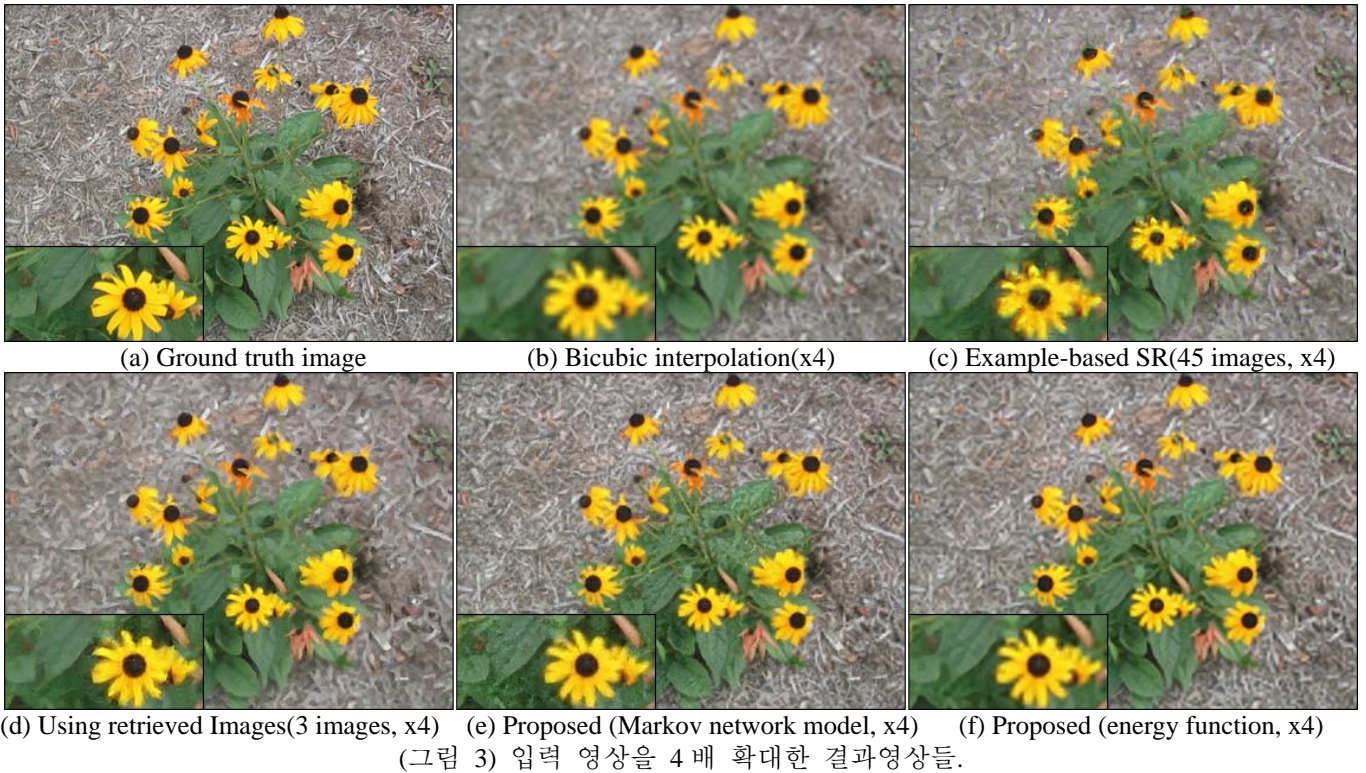
본 논문에서는 다양한 영상들을 최대한 활용할 수 있도록 영상검색과 예제기반 초해상도 방법을 통합하는 방법을 제안하였다. 외부 데이터베이스에 불필요한 정보들이 포함되어 초해상도 결과의 질을 저하시킨다는 것을 실험을 통해 보였고, 영상검색을 통해서 부분이 해결하였다. 또한 제안된 방법에서는 영상 내에 회전되어있는 정보들을 활용하기 위해서, 저해상도 패치 정합이 아니라 특징점을 기술한 기술자의 정합을 사용하였다. 합성단계에서는 학습단계에서 만들어 놓은 사전의 고해상도 패치를 입력영상의 특징점으로부터 얻을 수 있는 지역적인 방향으로 회전시켜 합성함으로써 회전 정보를 활용하였다. 결과적으로 영상검색을 통한 유사한 이미지 검색을 통해 불필요한 영상 정보들을 제거하고, 유사한 이미지들의 정보를 최대한 활용하여 좋은 결과를 얻었다. 이를 실험을 통해 검증하였다.

6. Acknowledgement

This research was supported by MKE (The Ministry of Knowledge Economy), Korea, under the Human Resources Development Program for Convergence Robot Specialists support program supervised by the NIPA (National IT Industry Promotion Agency).

참고문헌

- [1] W. Freeman, T. R. Jones, E. C. Pasztor. "Learning low-level vision", International Journal of Computer Vision (IJCV), 40(1), pp25-47, 2000.
- [2] W. Freeman, T. R. Jones, E. C. Pasztor. "Example-based SuperResolution", IEEE Comp. Graph. Appl, (2), p56-65, 2002.
- [3] D. Glasner, S. Bagon, M. Irani. "Super-Resolution from a Single Image" International Conference on Computer Vision (ICCV), p349-356, 2009.
- [4] G. Freedman, R. Fattal. "Image and Video Upscaling from Local Self-Examples" SIGGRAPH, 28(3), p1-10, 2011.
- [5] Z. Lin, H. Shum. "Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation", IEEE Transactions on Pattern Analysis and Machine Intelligence(PAMI), 26(1), p83-97, 2004.
- [6] J. Sun, J. Zhu, M. F. Tappen. "Context-Constrained Hallucination for Image Super-Resolution", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), p231-238, 2010.



- [7] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, C. Bray. "Visual Categorization with Bags of Keypoints", European Conference on Computer Vision(ECCV) workshop on Statistical Learning in Computer Vision, p1-22, 2004.
- [8] S. Lazebnik, C. Schmid, J. Ponce. "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), p2169-2178, 2006.
- [9] D. Nister, H. Stewenius, "Scalable Recognition with a Vocabulary Tree", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), p2161-2168, 2006.