

GÖRSEL BELİRGİNLİK GÜDÜMLÜ POZLANDIRMA BİRLEŞTİRİMİ

VISUAL SALIENCY GUIDED EXPOSURE FUSION

Kubra Mammadova, Erkut Erdem, Aykut Erdem

Bilgisayar Mühendisliği Bölümü

Hacettepe Üniversitesi

kubra.memmedli@gmail.com, {erkut, aykut}@cs.hacettepe.edu.tr

ÖZETÇE

Pozlandırma birleştirimi, bir sahnenin farklı pozlamalar altında yakalanan düşük dinamik aralığa sahip görüntülerinden yine düşük dinamik aralığa sahip ancak kontrast bakımından çok daha zengin bir görüntüsünü yaratmayı amaçlayan popüler bir çoklu pozlama birleştirme yöntemidir. Bu yaklaşımda birleştirme işlemi, kontrast, renk doygunluğu ve pozlandırma değeri gibi üç farklı yerel kalite ölçütü kullanılarak gerçekleştirilmektedir. Bu çalışmada amacımız, pozlandırma birleştirimi yönteminin görsel belirginliğe bağlı tanımlanan özgün bir kalite ölçütü ile iyileştirilmesidir. Bu yeni ölçüt sayesinde mevcut ölçütlerle yakalanamayan bir görüntüde görsel bakımdan öne çıkan alanlar yakalanarak bu belirgin alanlara daha çok önem verilmesi sağlanmaktadır. Yaptığımız deneyler, önerilen bu yeni belirginlik güdümlü yöntemin esas alınan pozlandırma birleştirimi yöntemine kıyasla genel olarak daha canlı çıktılar ürettiğini ve çıktı görüntülerindeki sınırların daha keskin olmasına yol açtığını göstermektedir.

ABSTRACT

Exposure Fusion is a popular multi-exposure image fusion method which blends a set of differently exposed low dynamic range images of a scene to obtain another low dynamic range but contrast rich image. This approach carries out the integration process by using three local quality measures, namely contrast, saturation and exposedness. Our aim in this study is to extend the exposure fusion method by incorporating a novel visual saliency based quality measure. This new measure captures the parts of the scene that grabs our attention and gives more prominence to these salient regions, which is otherwise impossible by the previous measures in use. Our experiments show that, as compared to the exposure fusion method, our saliency-guided approach gives more vivid results and leads to sharp boundaries in the output images.

1. GİRİŞ

Gerçek dünyadan elde edilen görüntüler mevcut sayısal kameraların kısıtlı dinamik aralığa sahip olmaları nedeniyle temsil ettikleri sahneleri ancak sınırlı olarak yakalarlar. Bu görüntüler özellikle düşük veya aşırı pozlandırma (exposure) altında yakalandıklarında ya çok az doymuş (under-saturated) ya da çok doymuş (over-saturated) alanlar içermekte ve bundan dolayı ilgili sahnedeki bazı detaylar kaybolmaktadır. Bu bakımdan geniş dinamik aralığa (high dynamic range - HDR) sahip sayısal görüntülerin elde edilmesi görüntü işlemede üzerinde çalışılan önemli problemlerden biridir [1].

Son yıllardaki gelişmeler sonucunda çeşitli özel donanımlar (örn. [2]) ve algoritmalar (örn. [3]) kullanılarak sayısal görüntülerin dinamik aralığının artırılabilmesi mümkün hale gelmiştir. Bu çözümler basitçe her piksel için

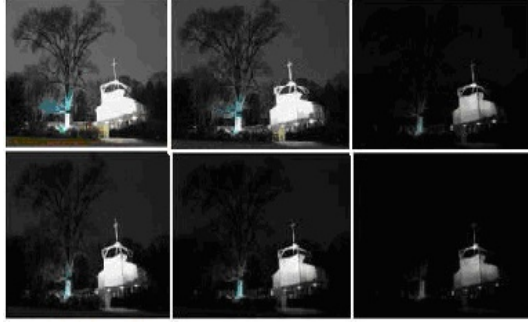
birbirinden farklı parlaklık değerlerinin yakalanması ve bunlar üzerinden yüksek bir dinamik aralığa sahip görüntünün hesaplanması prensibine dayanmaktadır. Bunu takiben yaratılan bu yüksek kontrastlı görüntülerin düşük dinamik aralığa sahip ortamlar (CRT, LCD ekranlar, vb.) için tekrardan oluşturulması başlı başına zorlu bir problem olarak karşımıza çıkmaktadır. Literatürde bu problemin çözümü için iki farklı yaklaşım önerilmiştir [4]: *Ton eşleme (tone mapping)* tabanlı [5-9] ve *pozlandırma birleştirmesi (exposure fusion)* baz alan [10-12] yöntemler.

Ton eşleme yaklaşımı, yüksek dinamik aralığa sahip görüntüler sınırlı dinamik aralığa sahip cihazlarda gösterilirken bunların görünüşlerin yaklaşık olarak korunması amacıyla bir renk kümesinin bir başka renk kümesine dönüştürülmesi olarak özetlenebilir. Ton eşleme operatörleri yerel ve global olmak üzere sınıflandırılabilir. Yerel ton eşleme operatörleri uzlamsal olarak değişen bir eşleme işlemi uygulayarak bazen doğallıktan uzak olmasına rağmen daha hoş bir görüntünün elde edilmesini sağlamaktadırlar. Global operatörleri ise görüntünün bütünü için tek bir yeniden eşleme uygulamaktadırlar. Bu nedenle yerel operatörlere kıyasla daha hızlı çalışmakta ama çoğunlukla daha kötü sonuçlar verebilmektedirler.

Ton eşlemeye bir alternatif olan pozlandırma birleştirmesi yaklaşımı [10] ise aynı sahnenin farklı pozlamalar altında yakalanan düşük dinamik aralığa sahip görüntülerinden yine düşük dinamik aralığa sahip ancak kontrast bakımından çok daha zengin bir görüntü yaratmaktadır. Bu yöntemin önemli avantajlarından biri görüntülerin elde edildiği pozlama sürelerinin ve kameranın tepki eğrisinin kalibrasyonunun bilinmesini gerektirmemesidir. Bu yaklaşım doygunluk değeri, kontrast değeri ve pozlama derecesi gibi basit yerel kalite ölçütlerine dayanmaktadır. Pozlandırma birleştirmesi, piramit ayrıştırmasına bağlı olarak çok ölçekli bir yapıda hesaplamaları gerçekleştirmektedir. Ton eşleme yaklaşımlarında olduğu gibi bu yaklaşımı izleyen çoğu çalışma da genelde birleştirilecek görüntülerin hizalanmış ve bundan dolayı uyumlu olduklarını varsaymaktadır. Şekil 1’de örnek bir pozlandırma birleştirimi işleminin sonucu gösterilmiştir.

Bu çalışmada, pozlandırma birleştirmesi yöntemi [10]; bir görsel belirginlik modeline dayanan ek bir yerel kalite ölçütü işlemlere dahil edilerek yeniden ele alınmıştır. Bu yeni ölçüt yukarıda sıralanan doygunluk değeri, kontrast değeri ve pozlama derecesi ölçütleriyle yakalanamayan bir görüntüde görsel bakımdan öne çıkan alanlara daha çok önem verilmesini sağlamak için önerilmiştir. Bu ölçüte bağlı iki farklı model üzerinden elde edilen sonuçlar görsel belirginliğin birleştirme işlemine dahil edilmesinin sonuçları iyileştirdiğini göstermektedir.

Bildirinin geri kalanı şu şekilde düzenlenmiştir: Bölüm 2’de pozlandırma birleştirimi yaklaşımı ve kullanılan görsel belirginlik modeli kısaca özetlenmiştir. Bölüm 3’te önerilen yeni kalite ölçütü ve birleştirme modelleri anlatılmıştır.



(a)



(b)

Şekil 1: Pozlandırma birleştirimi. (a) Pozlama dizisi, (b) Elde edilen zengin kontrastlı çıktı görüntüsü.

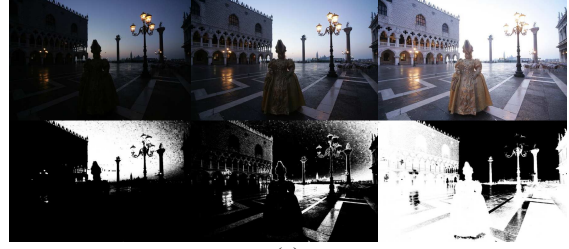
Bölüm 4'te yapılan deneysel incelemenin sonuçları verilmiştir. Bölüm 5'de bu sonuçlar özetlenmiş ve olası gelecek çalışmalara değinilmiştir.

2. İLGİLİ ÇALIŞMALAR

2.1. Pozlandırma Birleştirimi Modeli [10]

Sayısal görüntüleme cihazları ile elde edilen resimlerin hemen hepsi düşük veya aşırı pozlamaya maruz kalmalar nedeniyle düz, renksiz, aşırı ışıklanmış veya karanlıkta kalan ve detay içermeyen bölgeler içerir. Pozlandırma birleştirimi yöntemi [10], tekil olarak bakıldığında bu tarz alanlar içeren düşük dinamik aralığa sahip bir grup görüntü ($\{I_k\}_{k=1}^N$) üzerinden zengin kontrasta sahip görsel açıdan daha iyi bir görüntü (R) elde etmeyi amaçlamaktadır. Bu yaklaşımın işleyişi, çıktı görüntüsündeki bir pikselin değerinin diğer görüntülerde aynı konumda yer alan piksellerin ağırlıklı toplamı cinsinden tanımlanmasına dayanmaktadır. Bu şekilde, ilgili işlemin bu birleştirmede kullanılacak ağırlıkların belirlenmesi indirgendiği söylenebilir; öyle ki yukarıda değinilen problemler alanlara daha düşük ağırlık değerleri atanmakta ve onların birleştirmeye kısıtlı oranda dahil edilmesi sağlanmaktadır.

Bu bağlamda, Mertens ve diğerleri [10] şu üç ölçütün kullanılmasını önermişlerdir: *kontrast*, *renk doygunluğu* ve *pozlandırma değeri*. Kontrast ölçütü, doku ve kenarlar gibi önemli unsurlara yüksek ağırlıklı değerler atanmasını sağlarken, renk doygunluğu ise resimde kullanılan doymuş renklerin görüntünün daha canlı olmasını sağlayacağından bu tarz piksellere daha yüksek ağırlıklar atamaktadır. Pozlandırma ölçütü ise düşük ve aşırı pozlandırmaya sahip alanlara düşük ağırlık verilmesini zorlamaktadır. Sonuçta, bu değinilen ölçütlere göre girdi görüntülerindeki piksellerin ağırlıkları aşağıdaki denklemle hesaplanmaktadır:



(a)



(b)

Şekil 2: (a) Ağırlık haritalarıyla birlikte girdi resimleri, (b) birleştirilmiş görüntü ([10]'dan alınmıştır).

$$W_{ij,k} = (C_{ij,k})^{\omega_c} \times (S_{ij,k})^{\omega_s} \times (E_{ij,k})^{\omega_E} \quad (1)$$

Burada C , S , E sırasıyla kontrast, renk doygunluğu ve pozlandırma değerlerini, ω_c , ω_s , ω_E ilgili ölçütlerin önemini belirlemeye yarayan ağırlık parametrelerini ve (ij,k) , k görüntüsündeki (ij) 'nci pikseli ifade etmektedir. Eğer herhangi bir ağırlık parametresinin (ω) değeri 0 olarak alınırsa, bu o ölçütün birleştirmeye dahil edilmemesine karşılık gelmektedir.

Bu ağırlıklar kullanılarak N farklı görüntünün birleştirimi hesaplanan ağırlıkların $W_{ij,k}$ aşağıdaki şekilde normalize edilmesi ile

$$\hat{W}_{ij,k} = \left[\sum_{k'=1}^N W_{ij,k'} \right]^{-1} W_{ij,k} \quad (2)$$

aşağıdaki ağırlıklı toplam alma denklemiyle gerçekleştirilmektedir:

$$R_{ij} = \sum_{k=1}^N \hat{W}_{ij,k} I_{ij,k} \quad (3)$$

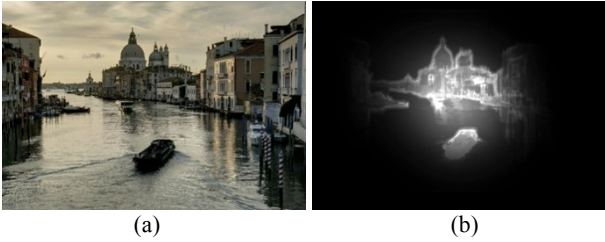
Şekil 2'de aynı sahnenin fazla pozlamalar ile yakalanan görüntüleri üzerinden pozlandırma birleştirimi kullanılarak elde edilen çıktı görüntüsü, girdi görüntülerinin ağırlık haritaları ile birlikte gösterilmiştir. Gözlemleneceği üzere sonuç görüntüsü girdi görüntülerine kıyasla daha zengin bir kontrasta sahipken sahnenin detaylarını da tek başına daha iyi korumaktadır.

2.2. Görsel Belirginlik Modeli [15]

Yakın tarihlerde görsel belirginlik hesabına berimsel yöntemler öneren çalışmalara olan ilgi ciddi miktarda artmış ve bu modeller bilgisayarlı görü ve görüntü işlemenin farklı problemlerinin çözümünde kullanılmaya başlanmıştır. Görsel

belirginlik, özetle bazı öğeleri komşularından ayıran ve bu sebeple biz insanların dikkati çeken farklı, nesnel algısal özellikler olarak tanımlanabilir. Itti ve diğerleri [13] tarafından önerilen, alt düzey ipuçlarının çok ölçekli bir yapıda incelendiği belirginlik modelinin ardından literatürde çok farklı modeller önerilmeye başlanmıştır. Bu modellerin yakın tarihli bir incelemesi için [14]'e başvurulabilir.

Bu çalışmada, pozlandırma birleştirmeni iyileştirmek amacıyla Margolin ve diğerleri tarafından önerilen yeni bir belirginlik modeli [15] kullanılmıştır. Bu model, görüntüde dikkat çeken ve bu manada görüntünün içeriğini en iyi ifade eden bölgeleri tespit ederken bir bölgenin yerel ve global bağlamı içinde ne kadar farklı olduğunu hesaplamaya çalışmaktadır. Bu nedenle, bu görsel belirginlik modeli, resimde sadece belirgin olan nesneleri değil, aynı zamanda arka planda olan ayırt edici nesneleri de belirgin olarak işaretlemeye imkan sağlamaktadır. Bu modeli, önceki görsel belirginlik modellerinden ayıran en önemli fark bu genel farklılık hesabını yaparken üç temel etkeni incelemesidir: 1. Desen farklılığı, 2. Renk benzersizliği ve 3. Görüntüleri ait önbilgiler. İncelenen ilk iki etken ayırt edici renk ve desen içeren alanlara yüksek belirginlik değeri atanmasını ve birbütün alanlara düşük belirginlik değeri verilmesi sağlamaktadır. Kullanılan önbilgiler ise doğal görüntülerden elde edilen bazı özellikleri (görüntü merkezinin önemli olması, belirgin piksellerin gruplanmış olması, vb.) kullanarak belirginlik hesabını daha da iyileştirilmesine neden olmaktadır. Şekil 3'te bu model kullanılarak hesaplanmış bir örnek belirginlik haritasına yer verilmiştir.



Şekil 3: (a) Girdi görüntüsü, (b) Görsel belirginlik haritası.

3. GÖRSEL BELİRGİNLİK GÜDÜMLÜ POZLANDIRMA BİRLEŞTİRİMİ

Bu bölümde pozlandırma birleştirmeni yaklaşımına görsel belirginlik modelini nasıl dahil ettiğimiz anlatılmıştır. Bölüm 2'de değinildiği üzere pozlandırma birleştirmeni yöntemi zengin kontrasta sahip çıktı görüntüsünü elde etmek için üç farklı yerel kalite ölçütü kullanılmakta bu bunlara bağlı olarak girdi görüntülerindeki her piksele farklı bir ağırlık değeri atamaktadır (bkz. Denklem (1)-(3)).

Görsel belirginliği işlemlere dahil etmek amacıyla önerilen ilk model farklı pozlama altında elde edilen her görüntüden çıkartılan belirginlik haritalarını dördüncü bir kalite ölçütü olarak eklemekte ve buna bağlı olarak yeni bir ağırlık hesabına gitmektedir:

$$W_{ij,k}^1 = (Z_{ij,k})^{\omega_Z} \times (C_{ij,k})^{\omega_C} \times (S_{ij,k})^{\omega_S} \times (E_{ij,k})^{\omega_E} \quad (4)$$

Burada $Z_{ij,k}$, k 'ninci girdi görüntüsündeki (i,j) 'nci pikselin belirginlik değerini ve ω_Z de ilgili ağırlık parametresini ifade etmektedir. Böylece ilgili bütünleşik yerel kalite değeri ek olarak ilgili görüntülerde dikkat çeken kısımlara daha fazla

ağırlık verilmesine ve bu manada resmin içeriği bakımından ilgi çekmeyen bölgelere daha düşük önem verilmesine neden olmaktadır.

Önerdiğimiz ikinci model ise aşağıdaki birleştirmeni kullanmaktadır:

$$W_{ij,k}^2 = \frac{(Z_{ij,k})^{\omega_Z} + (C_{ij,k})^{\omega_C}}{2} \times (S_{ij,k})^{\omega_S} \times (E_{ij,k})^{\omega_E} \quad (5)$$

Bu modeli, ilk modelden ayıran ana nokta belirginlikle kontrast derecesini birleştiren yeni bir yerel ölçütün bunların ortalaması alınarak tanımlanmış olmasıdır.

4. DENEYSEL SONUÇLAR

Deneylerde kullandığımız resimlerde, kalite ölçütlerine eşit öncelik verilmesi adına bütün ağırlık parametreleri 1 değerine sahip kabul edilmiştir ($\omega_Z = \omega_C = \omega_S = \omega_E = 1$). Önerilen görsel belirginlik güdümlü modellerin görsel iyileştirmelerini gözlemlemek için bu modellerle elde edilen sonuçlar orijinal pozlandırma birleştirmeni modelinin [10] sonuçları ile kıyaslanmıştır. Test ettiğimiz bazı girdi resimleri için elde ettiğimiz sonuçlar Şekil 4'te gösterilmiştir. En yukarıda iki ayrı sahnenin farklı pozlamalar altında yakalanan girdi görüntüleri, onun altındaki sıralarda da sırasıyla Pozlandırma birleştirmeni ve önerilen belirginlik güdümlü Model 1 ve Model 2'nin sonuçları verilmiştir. Şekil 5'te ise Şekil 4'te gösterilen sonuçlardan detay görüntülerine yer verilmiştir.

Mertens ve diğerleri, farklı pozlandırmalar arasındaki parlaklık değişikliğinin çok yüksek miktarlarda olması durumunda pozlandırma birleştirmeni modelinin ilgili çıktı görüntüsünde solgun renklere sebebiyet verebildiğini rapor etmişlerdir [10]. Bu bağlamda, önerdiğimiz her iki model ile elde edilen görüntülerin genelde orijinal pozlandırma birleştirmeni yönteminin sonucuna oranla daha canlı renkler vermektedir. Buna ek olarak, Model 1, diğerlerine göre resmin belirgin bölgelerinde çok daha keskin sonuçlar üretmektedir. Bildiride yer verdiğimiz örneklerin daha yüksek çözünürlüklü hallerine ve ek sonuçlara bu çalışma için hazırladığımız İnternet sayfasından erişilebilir¹.

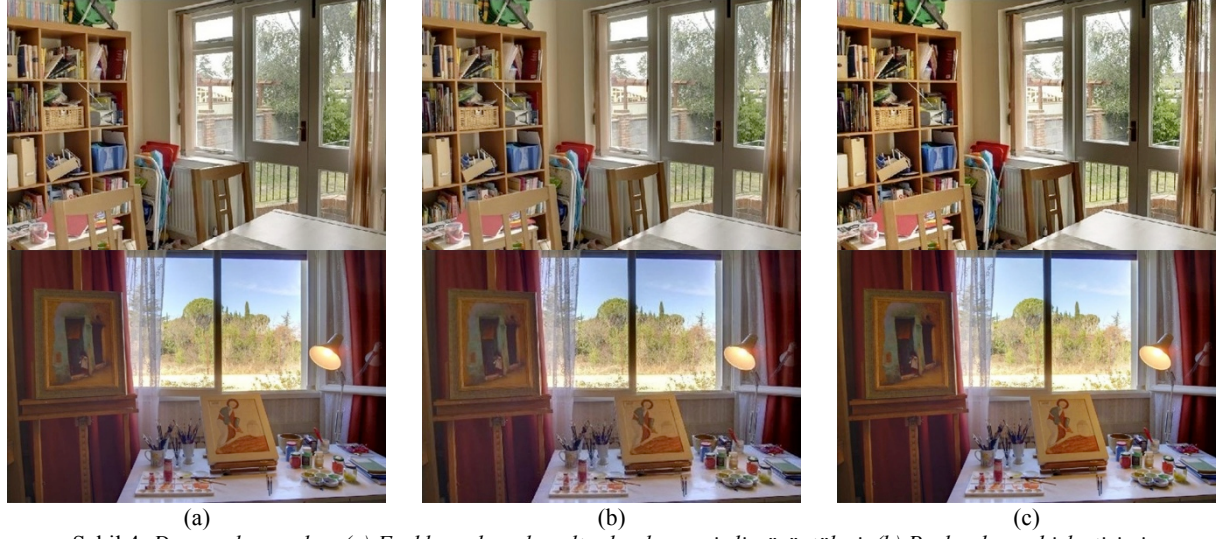
5. SONUÇ

Bu çalışmamızda farklı pozlamalar altında yakalanan girdi görüntülerini birleştirerek daha zengin kontrasta sahip bir çıktı görüntüsü üreten pozlandırma birleştirmeni [10] yaklaşımı, birleştirme işlemi için görsel belirginliğe ait yeni bir kalite ölçütü ile yeniden ele alınmıştır. Önerilen ölçüt üzerinden tanımlanan iki yeni model, esas alınan yönteme kıyasla daha canlı sonuçlar verirken çıktı görüntülerindeki kenarların daha keskin olmasını da sağlamaktadır. Elde edilen sonuçların başarısı yeni ölçütte kullanılan görsel belirginlik modelinin başarısı ile orantılıdır. Bu bakımdan bir sonraki aşamada literatürde yeni önerilmeye başlanan *belirgin nesne tespiti* yöntemlerinin [16] kullanımı araştırılacaktır.

6. TEŞEKKÜR

Bu çalışma kısmen TÜBİTAK 112E146 nolu proje tarafından desteklenmiştir.

¹ <http://dev.cs.hacettepe.edu.tr/~n11227207/SIU>



Şekil 4: Deneyisel sonuçlar. (a) Farklı pozlamalar altında alınan girdi görüntüleri, (b) Pozlandırma birleştirimi, (c) Belirginlik güdümlü birleştirim (Model 1), (d) Belirginlik güdümlü birleştirim (Model 2).



Şekil 5: Şekil 4'te gösterilen sonuçlardan detay görüntüleri. (a) Pozlandırma birleştirimi, (b) Model 1, (c) Model 2.

7. KAYNAKÇA

- [1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, "High dynamic range imaging: acquisition, display and image- based lighting. Second Edition," Morgan Kaufmann, San Francisco, 2010.
- [2] T. Mitsunaga, S.K. Nayar, "High dynamic range imaging: Spatially varying pixel exposures," In *CVPR* 2000.
- [3] P.E. Debevec, J. Malik, "Recovering high dynamic range radiance maps from photographs," In *SIGGRAPH* 1997.
- [4] A.O. Akyüz, M.L. Eksert, M.S. Aydin, "An Evaluation of Image Reproduction Algorithms for High Contrast Scenes on Large and Small Screen Display Devices," *Computers & Graphics*, Vol. 37, No. 7, pp. 885–895, November 2013.
- [5] F. Durand, J. Dorsey, "Fast Bilateral Filtering for the Display of High-Dynamic-Range Images," *ACM Trans. Graphics*, Vol. 21, No. 3, pp. 257-266, July 2002.
- [6] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photo-graphic tone reproduction for digital images," *ACM Trans. Graphics*, Vol. 21, No. 3, pp. 267-276, July 2002.
- [7] Y. Li, L. Sharan, E.H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *ACM Trans. Graphics*, Vol. 24, No. 3, pp. 836-844, July 2005.
- [8] R. Mantiuk, S. Daly, L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graphics*, Vol. 27, No. 3, pp. 68:1–10, August 2008.
- [9] Q. Shan, J. Jia, M. Brown, "Globally Optimized Linear Windowed Tone Mapping," *IEEE Trans. Vis. Comput. Graph.*, Vol. 16, No. 4, pp. 663-675, July-August 2010.
- [10] T. Mertens, J. Kautz, F. Van Reeth, "Exposure Fusion," In *Pacific Graphics* 2007.
- [11] A.A. Goshdasby, "Fusion of multi-exposure images," *Image and Vision Computing*, Vol. 23, No. 6, pp. 611–618, June 2005.
- [12] W. Zhang, W.-K. Cham, "Gradient-directed composition of multi-exposure images," In *CVPR* 2010.
- [13] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. and Mach. Intell.*, Vol. 20, No. 11, pp. 1254–1259, November 1998.
- [14] A. Borji, L. Itti, "State-of-the-Art in Visual Attention Modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 35, No. 1, pp. 185–207, January 2013.
- [15] R. Margolin, A. Tal, L. Zelnik-Manor, "What Makes a Patch Distinct?," In *CVPR* 2013.
- [16] A. Borji, D. N. Sihite, L. Itti, "Salient Object Detection: A Benchmark," In *ECCV* 2012.