Çok-odaklı Gürültülü Görüntülerin Birleştirilmesinde Kıstas Fonksiyonlarının Değerlendirilmesi

Evaluation of Criterion Functions for the Fusion of Multi-focus Noisy Images

Veysel Aslantas, Rifat Kurban

Bilgisayar Mühendisliği Bölümü Erciyes Üniversitesi
{aslantas,rkurban}@erciyes.edu.tr

Özetçe

Sayısal görüntüler, pek çok pratik uygulamada, algılayıcılar ve iletişim kanalları tarafından üretilen dürtü gürültüsü (DG) ile bozulurlar. Bu gürültü, çok-odaklı görüntü birleştirme uygulamalarında netliğin doğru bir şekilde tespit edilmesini engeller. Bu bildiride, geleneksel kıstas fonksiyonlarının ve frekans seçici ağırlık ortanca süzgeci (FSAO) tabanlı kıstas fonksiyonunun çok-odaklı görüntülerin gürültü altında birleştirilmesindeki başarımları değerlendirilmiştir. Deneysel sonuçlar, FSAO'nun diğer odaklama ölçütlerine göre daha iyi sonuçlar verdiğini ortaya koymuştur.

Abstract

Digital images are distorted by impulsive noise (IN) which caused by image sensors and communication channels in many practical applications. This noise may cause miscalculation of sharpness values in image fusion applications. In this paper, traditional criterion functions and frequency selective median filter (FSWM) based criterion function are evaluated for the fusion of multi-focus images in the presence of IN. Experimental results show that the FSWM filter is better than the other focus measures.

1. Giriş

Görüntü birleştirme, son yıllarda görüntü analizinde ve bilgisayarlı görü uygulamalarında önemli bir konu haline gelmiştir [1], [2]. Görüntü birleştirme, hem geometrik ve algısal parametreleri değişebilen bir algılayıcı, hem de çoklu-algılayıcılar kullanılarak elde edilen görüntülerin birleştirilmesiyle yeni ve gelişmiş bir görüntü üreten görüntü işleme tekniklerine denir. Bu yöntemle elde edilen birleştirilmiş görüntü, insan/makine algılaması ve bölümleme, öznitelik çıkartma, nesne tanıma gibi görüntü işleme görevleri için daha uygun bir hale gelir [3], [4].

Net bir şekilde odaklanmış olan görüntüler, doğal yapıları itibariyle bulanık görüntülerden daha çok bilgi ihtiva ederler. Bir sahnenin net bir görüntüsü, optik olarak diyafram açıklığı çapının küçük değerlere ayarlanmasıyla elde edilir. Ancak, bu durum bazı ciddi pratik problemlere sebep olur. [5], [6]. Diğer taraftan, her tarafı net görüntü, sahnenin bulanık görüntüsüne bir ters-evrişim (deconvolution) işlemi uygulanarak da elde edilebilir [7]. Ancak bu yaklaşım, kamera parametrelerin ve/veya bir yığın kamera kalibrasyon işlemleri gerektiren kameranın nokta dağılım fonksiyonunun bilinmesini gerektirir [4]. Bir sahnenin görüntüsü gerçek bir lens vasıtasıyla elde edilirse, lensten belirli bir uzaklıkta bulunan yüzey üzerindeki noktalar odaklanırken, diğer noktalar uzaklıklarına göre bulanıklaşır [8]. Bundan dolayı, her tarafı odaklanmış olan bir görüntü elde etmek için o sahnenin iki veya daha fazla görüntüsü birleştirilebilir.

Çok-ölçekli dönüşümler tabanlı görüntü birleştirme yöntemlerin temel fikri, her bir kaynak görüntüye çok-çözünürlüklü bir ayrışma uygulanması, tüm bu ayrışma sonuçlarının birleştirilerek birleşik bir sunum oluşturulması ve son olarak bir ters çok-çözünürlüklü dönüşüm uygulanarak birleştirilmiş görüntünün inşa edilmesi aşamalarını içerir [3]. Bu yaklaşımların iyi bilinen örnekleri Laplacian piramidi [9] ve ayrık dalgacık dönüşümüdür (ADD) [2], [10]. Ancak, ADD gibi çok-çözünürlüklü yaklaşımlar genellikle kaynak görüntülerdeki kaymalardan çok fazla etkilenirler. Bu problemin üstesinden gelebilmek için, kaynak görüntülerden resim uzayında (spatial domain) bloklar seçmeye dayalı piksel-seviyeli görüntü birleştirme yöntemleri önerilmiştir [3]. Bu tekniğin temel fikri, birleştirilmiş görüntüyü elde edebilmek amacıyla kaynak görüntülerden net olan görüntü bloklarının seçilmesidir. Bu işlemin otomatikleştirilmesi, görüntünün bir parçasına ne derecede odaklanıldığının netliğini ölçebilen bir analitik kıstas fonksiyonunu gerektirir. Görüntülerin netliğini ölçen pek çok kıstas fonksiyonu önerilmiştir [12], [13], [14]. Ancak bu yöntemler, kameraların CCD algılayıcıları tarafından üretilen gürültüleri önemsemezler. Dürtü gürültüsü, görüntü algılayıcıları ve iletişim kanallarındaki istenmeyen durumlardan dolayı meydana gelir [15]. Bu gürültü, netlik değerinin çoğu kez yanlış hesaplanmasına sebep verir. Bu ise, görüntü birleştirmenin performansını önemli ölçüde düşürür.

Bu bildiride, literatürdeki geleneksel kıstas fonksiyonlarının ve frekans seçici ağırlık ortanca süzgeci (FSOA) tabanlı kıstas fonksiyonunun çok-odaklı görüntülerin gürültü altında birleştirilmesinde kullanımı değerlendirilmiştir.

2. Kıstas Fonksiyonları

Bu bölümde, odaklanma kalitesini ölçen kıstas fonksiyonları açıklanmıştır. $f(i, j)$, bir görüntünün (i, j) koordinatlarındaki bir pikselin gri seviye yoğunluk değeri olsun:

2.1. Varyans (VAR)

$M \times N$ boyutunda bir görüntü bloğu için varyans, bir kıstas fonksiyonu olarak kullanılabilir ve şu şekilde hesaplanır:

$$VAR = \frac{1}{M \times N} \sum_i \sum_j (f(i,j) - \bar{f})^2 \qquad (1)$$

burada $\bar{f}$, ilgili görüntü bölgesinin gri seviye piksel değerlerinin ortalamasıdır:

2.2. Değiştirilmiş Laplacian Toplamı (DLT)

Laplacian operatörü yüksek frekansların belirlenmesinde kullanılabilir. Ancak yatay ve dikey yöndeki ikinci türevler farklı işaretlere sahiptirler ve birbirlerini yok edebilirler. DLT, klasik Laplacian'dan farklı olarak kısmi ikinci türevlerin kendileri toplanmak yerine mutlak değerlerini alarak toplar:

$$\begin{aligned}\nabla^2 f(i,j) &= \left|\frac{\partial^2 f(i,j)}{\partial i^2}\right| + \left|\frac{\partial^2 f(i,j)}{\partial j^2}\right| \\ &= \left|2f(i,j) - f(i-1,j) - f(i+1,j)\right| \\ &\quad + \left|2f(i,j) - f(i,j-1) - f(i,j+1)\right|\end{aligned} \tag{2}$$

Kıstas fonksiyonu ise şu şekilde verilir:

$$DLT = \sum_i \sum_j \nabla^2 f(i,j) \text{ ve } \nabla^2 f(i,j) \geq T \tag{3}$$

burada T eşik değeridir.

2.3. Uzaysal Resim Frekansı (UF)

Uzaysal resim frekansı (spatial frequency) [16], bir görüntüdeki genel aktiviteyi ölçer. Bu yöntemde, Sobel operatörlerini kullanmak yerine, eğimler (gradient) komşu pikseller arasındaki birinci derece farklar üzerinden hesaplanır. Eğimlerin karesi alınarak farklar daha fazla vurgulanır:

$$UF = \sqrt{C^2 + R^2} \tag{4}$$

burada C ve R satır ve sütun gradyentleridir:

$$\begin{aligned}C &= \left[\frac{1}{M \times N}\sum_i \sum_j \left[f(i,j) - f(i-1,j)\right]^2\right]^{1/2} \\ R &= \left[\frac{1}{M \times N}\sum_i \sum_j \left[f(i,j) - f(i,j-1)\right]^2\right]^{1/2}\end{aligned} \tag{5}$$

3. Frekans Seçici Ağırlık Ortanca Süzgeci

Geometrik analizler ve kırınım analizleri, bulanıklığın yüksek uzaysal frekansların azalması sonucu ortaya çıktığını göstermektedir. Bu iyi bilinen gerçek, bulanıklığın derecesinin yüksek uzaysal frekanslardaki enerji miktarı ile ilgili olduğunu belirtir. Neredeyse tüm kıstas fonksiyonları görüntüdeki yüksek frekans bilgisinin miktarı üzerine bina edilmiştir. İyi odaklanmış görüntüler daha keskin kenarlara sahip oldukları için, bulanık olanlara göre daha fazla yüksek frekans içeriğine sahip olduğu düşünülebilir. Yüksek frekans içeriği ise bir eğim kestirici ile elde edilebilir çünkü bir eğim kestirici aslında bir yüksek geçiren süzgeçtir. Eğim kestirici ve yüksek geçiren süzgeç arasındaki ilişki incelendiğinde, FSAO süzgecinin temel yapısı elde edilmiş olur [15]. FSAO süzgeci sadece görüntüdeki yüksek frekans bileşenlerine tepki göstermek ile kalmaz, aynı zamanda dürtü gürültüsünden de etkilenmez. Böylece, görüntünün keskinliğini daha hassas bir şekilde ölçmüş olur. Yüksek geçiren bir süzgecin karakteristikleri, doğrusal olmayan ağırlıklı ortanca süzgeç ile birlikte güçlendirilebilir. Bir ağırlıklı ortanca süzgeç <W; F> ile ifade edilsin, burada W = [w_1, w_2,…,w_m] ve F= [f_1, f_2,…,f_m] sırasıyla, tekrarlama vektörü ve ayrık zamanlı kayma giriş vektörüdür. F'in elemanları $f_1 \lhd f_2 \lhd \ldots \lhd f_m$ şeklinde sıralanır. Ağırlık ortanca süzgecinin çıkışı önce her bir f_i 'nin w_i kadar tekrar edilmesi ve sonrasında geliştirilen vektörün ortanca değerinin seçilmesiyle hesaplanır. Örneğin; <[1, 2, 3, 2, 1]; [-3, -2, 0, 1, 4]> şeklinde bir süzgeç, *ortanca*{$f(i$-3), 2◊$f(i$-2), 3◊$f(i)$, 2◊$f(i$+1), $f(i$+4)} yapısına denk gelmektedir. Burada ◊ tekrarlama operatörüdür, ör:, w◊f , f 'nin w kere tekrar edileceğini gösterir. o*rtanca* ise vektörün ortanca değerini ifade etmektedir. FSAO süzgecini oluşturmak için ağırlıklı ortanca süzgeçler doğrusal olarak birleştirilebilir:

$$H = \sum_i a_i < W_i, F_i > \tag{6}$$

Bir FSAO süzgeci uygun $a_i \in \mathrm{R}$, W_i ve F_i'lerin belirlenmesi ile elde edilir. [15]'de otomatik odaklama için uygun bir yapı türetilmiştir:

$$\begin{aligned}H &= \text{ortanca}\{f(n-1), f(n), f(n+1)\} \\ &\quad - \frac{1}{2}\text{ortanca}\{f(n-3), f(n-2), f(n-1)\} \\ &\quad - \frac{1}{2}\text{ortanca}\{f(n+1), f(n+2), f(n+3)\}\end{aligned} \tag{7}$$

Bu çalışmada, [15]'de önerilen kıstas fonksiyonu otomatik odaklama yerine görüntü birleştirmede netliği ölçmek amacıyla kullanılacaktır. H_h ve H_v, sırasıyla yatay ve dikey yönlerdeki FSAO sonuçları olsun. Kıstas fonksiyonu bu durumda şu şekilde tanımlanır:

$$FSAO = \sum_i \sum_j [H_h{}^2 + H_v{}^2] \tag{8}$$

4. Çok-odaklı Görüntü Birleştirme

Bu çalışmada, sadece iki görüntünün birleştirilmesi göz önüne alınmıştır ancak, yapı kolaylıkla ikiden fazla görüntü için genişletilebilir. Resim uzayında net görüntü bloklarının seçilmesine dayalı çok-odaklı görüntü birleştirme algoritmasının temel adımları şu şekildedir:

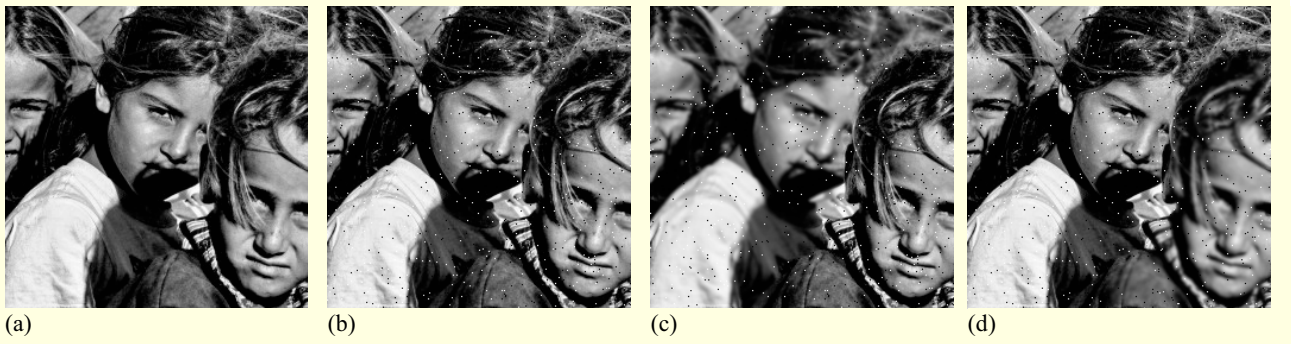
1. A ve B giriş görüntüleri üst üste gelmeyen ve aralarında boşluklar bulunmayan bloklara bölünür. A ve B'nin *i.* görüntü bloğu sırasıyla A_i ve B_i olarak adlandırılır.
2. A_i ve B_i için kıstas fonksiyonu vasıtasıyla netlik değerleri hesaplanır ve sonuçlar sırasıyla F_i^A ve F_i^B olarak adlandırılır.
3. Karşılıklı A_i ve B_i bloklarının netlik değerleri karşılaştırılarak daha net olan görüntü bloğu seçilir ve birleştirilmiş görüntü olan F_i çıkış görüntüsü şu şekilde elde edilir:

$$F_i = \begin{cases} A_i, & F_i^A > F_i^B \\ B_i, & diger \end{cases} \tag{9}$$

4. Birleştirilmiş görüntü üzerinde çoğunluk süzgeci uygulanır. Bu süzgecin amacı, izole blokları elimine etmektedir. Eğer bir blok A görüntüsünden kopyalandı ise ve çevresindeki blokların çoğunluğu B görüntüsünden ise, ilgili blok B'deki karşılığı ile değiştirilir. Benzer uygulama, tersi durum için de geçerlidir.

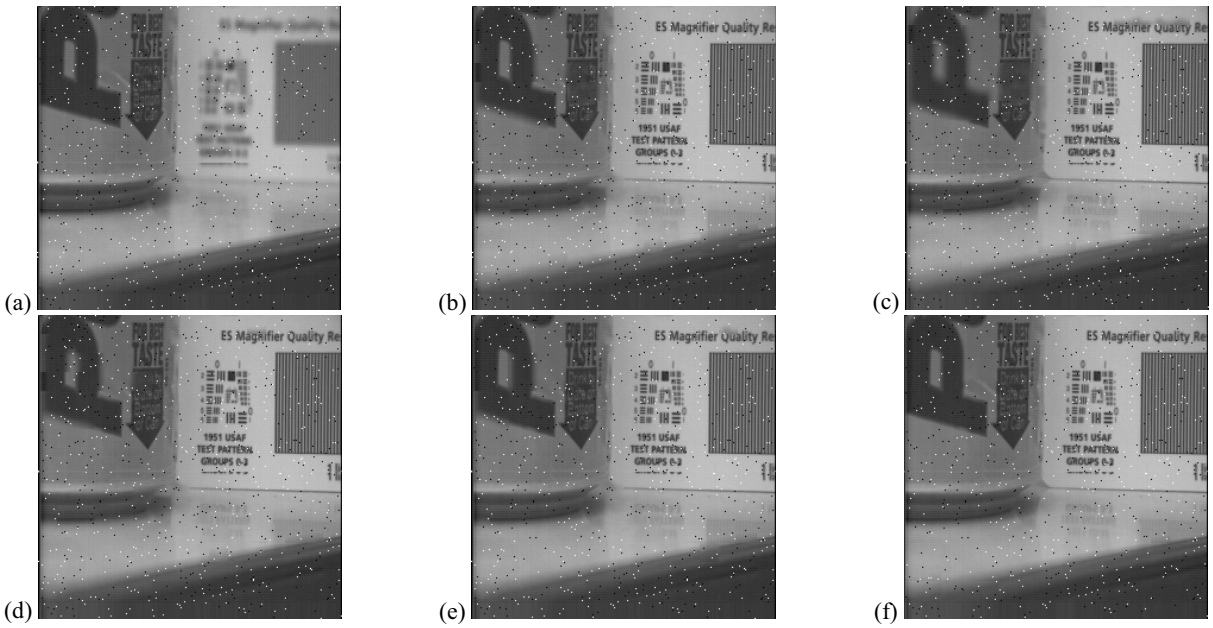
Tablo 1. *Çocuklar* görüntüsü için farklı pencere boyutu ve DG seviyesi altında 30 farklı koşmaya ait kıstas fonksiyonlarının MSE hatalarının ortalama ve standart sapmaları.

		0.5%			1%			1.5%			2%		
		8×8	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**	**8×8**	**16×16**	**32×32**
UF	Ort.	1,758	6,899	5,556	5,040	12,205	10,870	12,325	17,173	13,040	16,968	19,026	21,049
	Std.	0,709	2,391	2,741	1,167	3,856	6,846	2,961	5,782	6,953	3,569	6,731	9,175
VAR	Ort.	0,882	1,826	1,890	2,089	2,601	2,407	4,241	3,352	2,705	5,364	3,983	2,913
	Std.	0,340	0,628	0,653	0,633	1,125	0,994	1,110	1,210	0,906	1,139	1,687	1,311
DLT	Ort.	0,841	1,122	1,349	1,646	2,170	1,629	3,573	2,331	1,966	4,720	3,928	2,476
	Std.	0,366	0,545	0,126	0,489	1,304	0,666	1,241	0,820	0,759	1,448	1,348	1,130
FSAO	Ort.	*0,107*	*0,829*	*1,247*	*0,159*	*0,975*	*1,432*	*0,223*	*1,129*	*1,715*	*0,309*	*1,278*	*1,800*
	Std.	*0,024*	*0,165*	*0,103*	*0,090*	*0,278*	*0,391*	*0,114*	*0,364*	*0,337*	*0,155*	*0,506*	*0,531*



(a) (b) (c) (d)

Şekil 1. *Çocuklar* deney görüntüsü: (a) referans görüntü, (b) FSAO ile birleştirilmiş görüntü (blok boyutu=16×16), (c) bulanık kaynak görüntü (odaklama sağda), (d) bulanık kaynak görüntü (odaklama solda).



(a) (b) (c) (d) (e) (f)

Şekil 2. *Pepsi* deney görüntüsü ve 1% DG ile birleştirme sonuçları: (a) odak solda (b) odak sağda, (c) UF, (d) VAR, (e) DLT, (f) FSAO

5. Deneyler

Bu bölümde, yapay olarak bulanıklaştırılmış ve doğal çok-odaklı görüntüler üzerinde kıstas fonksiyonlarının performansını ölçmeye yönelik gerçekleştirilen deneyler verilmiştir. İlk deney her tarafı odaklı olan veya olduğu kabul edilen 256×256 boyutlardaki 8-bit'lik *Çocuklar* görüntüsü üzerinde yapılmıştır. *Çocuklar* görüntüsünün (Şekil 4(a)) bulanık versiyonları, görüntünün sol ve sağ yarılarının Gaussian süzgeç (r=1) ile bulanıklaştırılması ve farklı oranlarda rastgele DG ile bozulmasıyla elde edilmiştir. Örnek olarak, Şekil 4. (c) ve (d)'de 1% DG ile bozulmuş kaynak görüntüleri görülebilir.

Yapılan deneylerde kullanılan performans fonksiyonu, literatürde sıkça kullanılan referans tabanlı bir metrik olan ortalama karesel hatadır (MSE):

$$MSE = \frac{1}{M \times N} \sum_{i}^{M} \sum_{j}^{N} \left(R(i,j) - F(i,j) \right)^2 \qquad (10)$$

burada $R(i, j)$ ideal referans, $F(i, j)$ elde edilen birleşik görüntü, ve M, N ise görüntünün boyutlarıdır. Küçük MSE değerleri, daha iyi birleştirme performansı demektir.

Kıstas fonksiyonlarının performanslarını ortaya koymak için deneyler, çeşitli blok boyutları ve farklı seviyelerde DG ile yapılmıştır. Simülasyonlar, yapay DG gürültüsünün rastgele doğasından dolayı 30 kez gerçekleştirilmiştir. Karşılaştırma için Tablo 1, Şekil 1(c) ve (d)'nin birleştirilmesi ile ilgili bağımsız 30 farklı koşmanın ortalama MSE değerlerini ve standart sapmalarını vermektedir. Tablo 1'den görülebileceği üzere FSAO farklı gürültü seviyelerinde bile diğer fonksiyonlardan daha iyi sonuç vermektedir. Bütün deneyler göstermiştir ki, FSAO'nin MSE değerleri diğer yöntemlerden daha düşüktür. Buna ek olarak, hesaplanan standart sapma değerlerine göre de daha iyi olduğu görülmektedir. Bunlar, FSAO kıstasının diğerlerine göre daha güvenilir, sağlam ve kararlı olduğunu göstermektedir.

İkinci deney gerçek lens ile elde edilen görüntüler üzerinde gerçekleştirilmiştir (*Pepsi*). Görüntüler, kameradan farklı uzaklıkta birden fazla nesne içermektedirler. Deneylerde 1.5% oranında yapay dürtü gürültüsü kaynak görüntülere eklenmiştir. Kıstas fonksiyonları, Şekil 2.(a) (odak solda) ve Şekil 2.(b) (odak sağda) üzerinde 16×16 blok boyutları ile uygulanmıştır. Elde edilen birleşik görüntüler Şekil 2.(c)-(f)'de görülmektedir. Bu türden görüntüler için, birleştirme sonucuna ilişkin niceliksel bir değerlendirme gerçekleştirmek referans temelli metrikler ile mümkün değildir. Çünkü her tarafı odaklı olan bir kaynak görüntü mevcut değildir. Bundan dolayı, sonuçlar sübjektif olarak değerlendirilebilmektedir. Şekil 2(c)-(f)'de verilen birleşik görüntülerde pek çok problemli bölge görülmektedir. Örneğin, sağ üst köşelerdeki metin kısımları bulanıktır. Buna ek olarak, Pepsi tenekesi üzerinde ve masanın kenar kısmında yanlış kaynak görüntüden kopyalanmış bloklar vardır. Diğer taraftan, FSAO kullanılarak elde edilen birleşik görüntü, Şekil 2 (f), diğer tekniklerden daha iyi sonuç vermiştir.

6. Sonuçlar

Neredeyse tüm kıstas fonksiyonları, görüntü yüksek frekans bilgisi ile direkt ilişkilidir. Ancak, görüntüdeki gürültünün çoğunun da yüksek frekanslar ile ilişkili olduğu unutulmamalıdır. Bu problem, görüntü bloklarında netliğin yanlış olarak hesaplanmasına yol açar. Bu yüzden, netlik hesaplamalarında gürültünün etkisi olabildiğince azaltılmalıdır.

Görüntü birleştirmeden önce yapılan gürültü giderme girişimleri görüntülerin orijinal piksel değerlerini değiştirebileceğinden dolayı birleştirme algoritmalarının performansını da etkileyebilir. Bu durumda, birleştirme işlemi orijinal görüntüler üzerinde değil işlenmiş görüntüler üzerinde yapılmış olur. Bundan dolayı, görüntü birleştirme algoritmaları birleştirme aşamasından önce süzgeçleme operasyonları kullanmamalıdır.

FSAO kıstası, gürültüden etkilenmemesiyle birlikte ilgili görüntü bloğunun netliğini hassas bir şekilde hesaplayabilmektedir. Eşik seviye belirleme zorunluluğu olmadığından dolayı sezgisel seçimler yapılmak durumunda kalınmaz. Deneyler, farklı blok boyutları ve farklı seviyelerde DG için gerçekleştirilmiştir. Elde edilen sonuçlar, gürültülü görüntülerin birleştirilmesinde FSAO'nin diğer kıstas fonksiyonlarından niceliksel ve niteliksel olarak daha iyi olduğunu ortaya koymuştur.

7. Kaynakça

[1] Varshney, P.K., "Multisensor data fusion," *Electronics & Comm. Eng. Journal*, 9(6), 1997, 245-253.

[2] Zhang, Z., Blum, R.S., "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proc. of the IEEE*, 87, 1999, 1315-1326.

[3] Li, S., Kwok, J.T., Wang, Y., "Combination of images with diverse focuses using the spatial frequency," *Information Fusion*, 2, 2001, 169-176.

[4] Li, S., Kwok, J.T., Tsang, I.W., Wang, Y., "Fusing images different focuses using support vector machines," *IEEE Trans. on Neural Networks*, 16, 2004, 1555-1561.

[5] Pham, D.T., Aslantas, V., "Depth from defocusing using a neural network," *Pattern Recog.*, 32, 1999, 715-727.

[6] Aslantas, V., Pham, D.T., "Depth from Automatic Defocusing", *Optics Express*, 15, 2007, 1011-1023

[7] Subbaro, M., Wei, T.C., Surya, G., "Focused image recovery from two defocused images recorded with different camera settings," *IEEE Trans. on Image Processing*, 4, 1995, 1613-1628.

[8] Aslantas, V., "A depth estimation algorithm with a single image," *Optics Express*, 15, 2007, 5024-5029.

[9] Burt, P.T., Andelson, E.H, "The Laplacian pyraind as a compact image code," *IEEE Trans. on Communications*, 31, 1983, 532-540.

[10] Yocky, D.A., "Image merging and data fusion by means of the discrete two-dimensional wavelet transform," *J. of The Optical Society of America A-Optics Image Science and Vision*, 12 (9), 1995, 1834-1841.

[11] Huang, W. Jing, Z., "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognition Letters*, 28, 2007, 493-500.

[12] Aslantas, V., "Criterion functions for automatic focusing," in *10. Turkish Symp. on Artificial Intelligence and Neural Net.*, (Gazimagusa, TRNC), 2001, 301-311.

[13] Krotkov, E.P., "Focusing," *Int. Journal of Computer Vision*, 1 (3), 1987, 223-237.

[14] Groen, F.C.A., Young, I.T., Ligthart, G.,"A Comparison of Different Focus Functions for Use in Autofocus Algorithms," *Cytometry*,6, 1985, 81-91.

[15] Choi, K., Lee, J., Ko, S., "New autofocusing technique using the frequency selective weighted median filter for video cameras," *IEEE Trans. on Consumer Electronics*, 45, 1999, 820-827.

[16] Eskicioglu, A.M., Fisher, P.S., "Image quality measures and their performance",. *IEEE T. Commun.* 1995, 43, 2959-2965.