

SIFT Metodu ile Hedef Takibi

Nazım ÖZGEN¹, Müzeyyen SARITAŞ²

¹ Hava Kuvvetleri Komutanlığı, Çankaya, ANKARA

nzmzgn@gmail.com

² Gazi Üniversitesi, Elektrik-Elektronik Mühendisliği Böl., Maltepe-ANKARA

muzeyyen@gazi.edu.tr

Özet

Hedef takibinde en çok kullanılan yöntemler özellikli nokta tanımlayıcıları (Moravec, Harris, KLT ve SIFT), arka plan çıkartımı ve bölümlendirme yöntemleridir. Bu yöntemlerin kendilerine göre avantaj ve dezavantajları bulunmaktadır. Bu çalışmada, sabit/değişken arka planda, sabit/hareketli hedefi en iyi takip edebilecek yöntem olan SIFT metodu seçilmiştir. MATLAB ortamında hazırlanmış olan SIFT programı hedef takibinde kullanılmıştır. SIFT parametrelerinin hedef takibine etkileri incelenmiş ve SIFT parametrelerinin optimum değerleri elde edilmiştir. Bu optimum değerleri literatür ile uyumlu bulunmuştur.

Ayrıca, bu çalışmada, birbirini takip eden resim kareleri arasındaki hatalı eşleşmeleri yok etmek için ek bir program hazırlanmıştır. Bu program, optimum olmayan SIFT parametre değerlerinde kullanıldığında, hedef takibinde görülen sapmalar azalmıştır.

Resim çözünürlüğü ve bilgisayar işlemci hızı değiştirilerek hedef takibine etkileri incelenmiştir. Resim çözünürlüğü artırıldığında, SIFT kilit nokta sayısının, SIFT kilit noktalarının tespit süresinin ve hedefi bulma süresinin exponansiyel olarak arttığı gözlenmiştir. Bilgisayar işlemci hızı artırıldığında ise SIFT kilit noktalarını bulma süresinin ve hedefi bulma süresinin azaldığı belirlenmiştir.

Anahtar Kelimeler: Bilgisayar Destekli, SIFT Metodu, Hedef Takibi.

Target Tracking Using SIFT Method

Abstract

The most commonly used methods for tracking are point detectors (Moravec, Harris, KLT and SIFT), background subtraction and segmentation methods. Each method has its advantages and disadvantages with respect to each other. In this study, we choose SIFT method which is the best static/mobile target tracking method in the constant/variable background. SIFT program implemented in MATLAB was used in target tracking. The effects of SIFT parameters in target tracking were analyzed and optimum values of these parameters were obtained. These optimum values were found compatible with the literature.

Moreover, additional program has been written to minimize the probable false matchings between the picture frames. When this program was employed without optimum SIFT parameters, the deflection seen in target tracking was decreased.

The target tracking was analyzed by varying picture resolution and the speed of computer processor. When picture resolution was increased; it was seen that the number of SIFT keypoints, the duration of finding SIFT keypoints and the target were increased exponentially. When the speed of computer processor was increased, the time required for finding the SIFT keypoints and the target were decreased.

1. Giriş

En basit tanımıyla takip; hedefin konumu ardışık resimler düzleminde değişirken, hedefin güzergahını izlemek olarak tanımlanabilir. Diğer bir deyişle; takipçi, farklı video karelerinde izlenecek olan hedefe, değişmeyen etiketler atamalıdır.

Hedef takibi, takip edilecek hedefin bulunduğu ortamlar, hedefin alabileceği konumlar ve görüntünün alındığı kameranın konumu açısından farklılıklar göstermektedir. Burada hedef; az-orta-çok şiddetteki gürültüde, kendisiyle kıyaslandığında etrafında az-orta-çok sayıda karışık-net hedefler arasında bulunabilir. Hedefler üç boyutlu olmalarına karşın, alınan görüntüler iki boyutlu sınırlı olduğundan hedef, sonraki resim karelerinde boyutsal, döngüsel değişimlere maruz kalabilir. Ayrıca, hem hedeften hem de kameradan kaynaklanabilecek hareket ile, hedefin resim içindeki konumu da sürekli değişebilmektedir.

Yukarıda bahsedilen olumsuz faktörlerin etkisi, farklı uygulamalarda değişik seviyelerde gözlenmektedir. Bu çalışmada, bütün bu değişimler göz ardı edilmeden hedef takibi yapılması amaçlanmıştır.

2. Hedef Takip Yöntemleri

Hedef takibi konusunda literatürde karşılaşılan yöntemler [1] aşağıda kısaca açıklanmıştır.

Arka plan çıkartımı

Bu yöntem, sabit kamera ile sabit ortamda hareket eden nesnelerin tespitinde ve takibinde çok kullanışlı bir yöntemdir. Ayrıca bu yöntem, arka plan modeli olarak tanımlanan görünümün temsil edildiği ve bir sonraki resimde bu arka planda oluşan sapmaların tespit edildiği bir yöntemdir. Tanımlanan arka planda önemli bir değişim olması, hareketli bir nesnenin varlığına işaret etmektedir.

Bu yöntem, 1970'lerin sonunda çalışılmaya başlanmakla birlikte ancak 1997'de Wren'in yaptığı çalışma ile ünlenmiştir.

Wren'in çalışmasına göre, sabit arka plandaki her pikselin renk değeri, $I(x,y)$, üç boyutlu tek bir Gauss fonksiyonu ile modellenir. İlk resim karesindeki her (x,y) piksel için arka plana ait model çıkartıldıktan sonra, arka plana ait oluşturulan modeldeki sapmalar ön plan pikselleri olarak tanımlanır.

Bu yöntem işlem yükü açısından etkindir. Bu metodun en önemli kısıtlaması, sabit bir arka plan gerektirmesidir. Kameranın hareketli olması, arka planın bozulmasına sebep olmaktadır. Bu nedenle, sabit kamera uygulamalarında etkindir. Hareketli kamera uygulamaları da bulunmaktadır. Ancak, bu çözümler, küçük hareketlerin olduğu ve yüzeyel görünüm varsayımları altında oluşturulmaktadır.

Bölümleme

Burada temel amaç, resimde aynı özelliklere sahip bölgeleri tespit etmektir. Bir çok yöntem bulunmaktadır.

Bölümleme algoritmalarındaki amaç, resmi benzer alanlara bölmektir. Her bölümleme algoritması, iki temel probleme çözüm bulmalıdır. Bunlar, iyi bir bölümleme için uygun kriter ve etkin bir bölümleme için uygun bir yöntemdir [2].

Bu yöntem genellikle geniş bir örnek yelpazesi gerektirir. Bu yüzden iki veya daha fazla farkı ayırt edici özelliğin kullanıldığı, küçük örnek kümelerine sahip sınıflandırıcılara kullanılır.

Özellikli Nokta Tanımlayıcıları

Özellikli noktalarda aranan nitelikler [1], aşağıdaki şekilde tanımlanabilir. Buna göre, özellikli bir nokta;

- Açık, belli, tercihen matematiksel olarak iyi tanımlanmış olmalıdır.
- Resim düzleminde iyi tanımlanmış konuma sahip olmalıdır.
- Bu noktanın etrafındaki resim yapısı bilgi içeriği yönünden zengin olmalıdır.
- Resimde oluşabilecek değişimlere (boyutsal, döngüsel, parlaklık, görünüş) karşı kararlı, başka bir deyişle yüksek tekrar edilebilirliğe sahip olmalıdır.

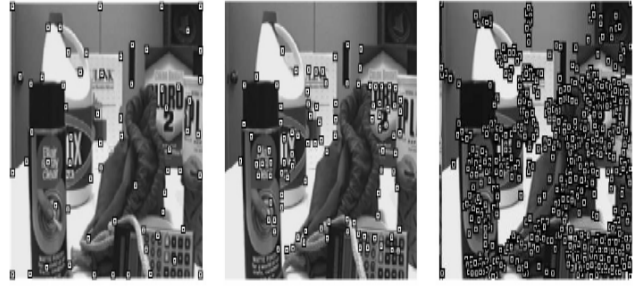
Literatürde var olan en önemli özellikli nokta bulan, nokta tanımlayıcıları; Moravec operatörü [3], Harris Algılayıcısı [4], KLT dedektörü [5], ve SIFT dedektörüdür [6] (Şekil 1). SIFT dedektörü kullanılarak belirlenen nokta tanımlayıcılarının; görüntüdeki döngüsel, boyutsal, üç boyutlu ve parlaklık değişimleri ile gürültüye karşı en dayanıklı yöntem olması sebebiyle bu çalışmada SIFT dedektörü kullanılmıştır.

3. SIFT Algoritması ve Bulgular [1,6]

SIFT Algoritması

Bu çalışmada kullanılan SIFT algoritması dört aşamadan oluşmaktadır:

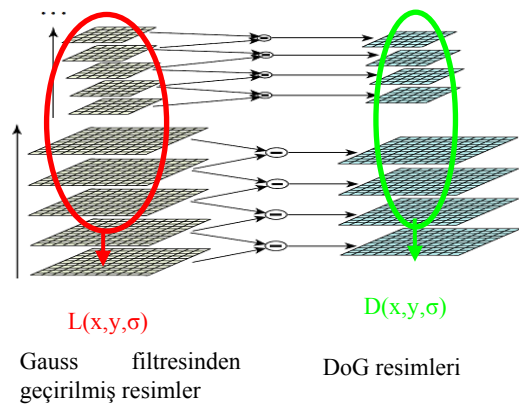
- Ölçeksel uzaydaki uç noktaların (minimum - maksimum) elde edilmesi
- Kilit noktaların konumlarının belirlenmesi
- Döngüsel değişime karşı dayanıklılık kazanılması
- Kilit nokta tanımlayıcıların bulunması



Şekil 1 : Özellikli nokta tanımlayıcılarının buldukları noktalar
a) Harris algılayıcısı b) KLT c) SIFT

Ölçeksel uzaydaki uç noktaların (minimum - maksimum) elde edilmesi

Ölçeksel uzayda sabit, değişmeyen kilit noktalar bulmak için birçok teknik kullanılabileceği düşünülmektedir. Bunlardan biri DoG (Difference of Gaussians) metoduydu. Koenderink (1984) ve Lindeberg (1994); yaptıkları çeşitli varsayımlar sonucunda ölçeksel uzayın (Şekil 2) Gauss fonksiyonu kullanılarak oluşturulabileceğini tespit etmişlerdir. SIFT algoritmasında da, DoG metodu kullanılmıştır.



Şekil 2 : Ölçeksel uzayın oluşumu.

Önce, ölçeksel uzay oluşturulur. Daha sonra kilit nokta olabilecek noktaların tespiti yapılır. Orijinal resim, farklı standart sapmaya sahip Gauss filtrelerinden geçirilir.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

I: orijinal resim, G: değişken orantıya sahip Gauss fonksiyonu olmak üzere, ölçeksel uzay fonksiyonu şu şekilde tanımlanır:

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \quad (2)$$

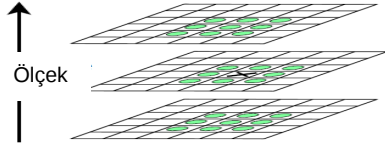
DoG (Diffrence of Gaussians) fonksiyonu olan $D(x, y, \sigma)$, biri diğerine göre k kadar orantılı büyüklüğe sahip Gauss filtreli iki resim arasındaki farkın sonucunda elde edilir:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

$\sigma \nabla^2 G = \partial G / \partial \sigma \approx [G(x, y, k\sigma) - G(x, y, \sigma)] / (k\sigma - \sigma)$ ve dolayısıyla $G(x, y, k\sigma) - G(x, y, \sigma) = (k-1) \cdot \sigma \nabla^2 G$ elde edilir.

Bu eşitlik, resmi Gauss filtresinden geçirmek ile Laplace dönüşümünü elde etmek arasında ciddi bir benzerlik olduğunu göstermektedir.

Burada, 26 adet nokta içinden sadece bir noktanın komşularına göre (Şekil 3) minimum veya maksimum olup olmadığı kontrol edilmektedir.



Şekil 3: Uç noktaların bulunması

Kilit noktaların konumlarının belirlenmesi[1]

Bir önceki aşamada, kararlı olmamasına rağmen tespit edilen bir çok kilit nokta adayı olmuştur. Bu aşamada kararlı olmayan yani düşük kontrasta sahip (dolayısıyla gürültüden daha çok etkilenen) veya kenarlarda zayıf olarak tespit edilmiş kilit noktaların ayıklanması sağlanır.

Düşük kontrasta sahip noktaların ayıklanması, DoG fonksiyonunun ($D(x,y,\sigma)$) kilit aday nokta baz olmak üzere ikinci dereceden Taylor serisi açılımı yapılarak gerçekleştirilir ve aşağıdaki eşitlik elde edilir:

$$\text{Uç noktaların yeni konumları, } \hat{x} = -\frac{\partial^2 D}{\partial x^2}^{-1} \frac{\partial D}{\partial x} \text{ ile}$$

$$\text{hesaplanır ve } D(\hat{x}) = D + \frac{1}{2} \frac{\partial^2 D}{\partial x^2} \hat{x}^T \text{ elde edilir.}$$

Burada, DoG fonksiyonu olan $D(x)$ ve türevi, her uç nokta için hesaplanır ve $|D(x)| < 0.02$ ise elenir.

Kenarlarda tespit edilen kilit noktalar ise aşağıdaki kriterlere göre elenir.

D : DoG fonksiyonu,
Heissan matrisi $H = [D_{xx}, D_{xy}, D_{yx}, D_{yy}]$,

a: büyük olan öz değer, b: küçük olan öz değer[1],

$r=a/b$ olmak üzere,

$R = (r+1)^2/r \Rightarrow$ SIFT için gerekli olan kriterdir

Döngüsel değişime karşı dayanıklılık kazanılması[1]

Algoritmanın bu aşaması, nesnenin iki boyutlu eksen etrafında dönmesine karşı dayanıklılık kazanılması için önemlidir. Bu aşamada her kilit nokta için bir veya birden fazla döngüsel atama yapılmaktadır.

Daha önce σ ölçeğindeki Gauss filtresinden geçirilmiş resim olan $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$ kullanılarak gradyent büyüklüğü $m(x,y)$ ve açısı $\theta(x,y)$ hesaplanır :

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1} \frac{(L(x+1, y) - L(x-1, y))}{(L(x, y+1) - L(x, y-1))}$$

Büyüklük ve yön hesaplaması, Gauss filtresinden geçirilmiş resimde bulunan kilit noktaların, komşu her pikseli için yapılır.

Kilit noktaların etrafında her biri 10 derecelik bir yönü kapsayan 36'lık bir döngüsel histogram oluşturulur. Histograma eklenen komşu her örnek nokta, kendi gradyent büyüklüğü ile ağırlıklandırılır.

Histogramda en büyük değere sahip nokta, } Kilit Nokta
%80'i içinde kalan noktadır } Tanımlaması

Noktaların sadece yaklaşık %15'i için çoklu döngüsel atama yapılmış olmasına rağmen bu işlem eşleşmelerdeki kararlılığı önemli ölçüde arttırmaktadır.

Kilit noktalara ait tanımlayıcıların tespit edilmesi[1]

Bu aşamada ise, diğer değişimlere (parlaklık değişimleri, üç boyutlu görünüm,vb.) karşı dayanıklılık kazanılır.

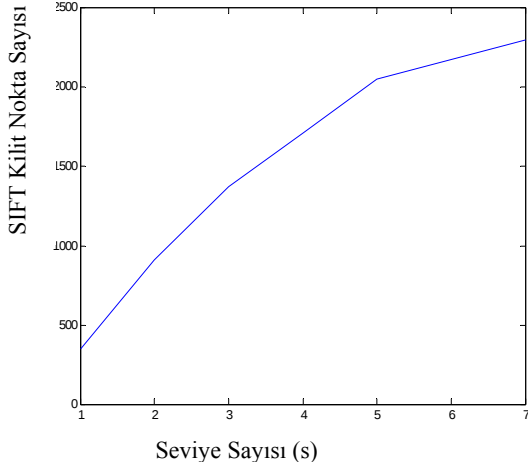
Önce, kilit nokta etrafında gradyent büyüklükleri ve açıları kilit noktanın ölçeksel büyüklüğü kullanılarak $n \times n$ 'lik alanda örneklenir. Döngüsel değişmezliği sağlamak için kilit nokta tanımlayıcının konumu ve gradyent açıları, kilit nokta açısı baz alınarak döndürülür.

Bu aşamada yapılanlar, döngüsel dayanıklılık kazanım işleminde yapılanlara çok benzemektedir. Sekizlik vektörler, kilit noktanın etrafında bulunan $n \times n$ 'lik alanın içinde kalan her bir nokta için tanımlanır. Dolayısıyla her bir kilit nokta için tanımlanan vektörün uzunluğu $n \times n \times r$ olur.

3. SIFT Parametrelerinin Hedef Takibine Etkileri[1]

Bu çalışmada, SIFT parametrelerinin ($s, |D(x)|, R, n \times n, r$) etkileri, 320x240 çözünürlüğe sahip resimler üzerinde 1.6 G Hz.lik işlemci kullanılarak incelenmiştir.

Ölçeksel uzaydaki her kademede bulunacak toplam seviye sayısı s değeri arttıkça, tespit edilen SIFT kilit nokta sayısı ve işlem süresi de artmaktadır. Yapılan testler sonucunda, tespit edilen SIFT kilit nokta sayısının, s değeri ile değişimi Şekil 4'te verilmiştir. Farklı s değerleri ($s=1-4$ arası) için hedef takibi yapılmış, Şekil 5'de görüldüğü gibi, s 'nin optimum değerinin 3 olduğu tespit edilmiştir.



Şekil 4: S değerinin SIFT kilit nokta sayısına etkisi



Şekil 5: Optimum $s=3$ değerinde hedef takibi

Bu çalışmada, düşük kontrasta sahip noktaların ayıklanması için gerekli olan eşik değerinin ($D(x)=0.002-2$ arası), hedef takibine etkisi araştırılmıştır. Optimum hedef takibi, $D(x)=0.02$ değerinde gerçekleşmiştir. (Bkz. Şekil 6)

Kenarlarda tespit edilen noktaların ayıklanması için gerekli optimum eşik değeri, R 'nin ($R=2-17$ arasında) hedef takibine etkisi incelenmiş, Şekil 7'de da görüldüğü gibi, 10 olarak bulunmuştur.

Toplam kilit nokta sayısı, örnekleme alanı ($n \times n$) ile ters orantılı olarak, işlem süresiyle ise doğru orantılı olarak değişmektedir. Örnekleme alanının farklı değerleri ($n=2-6$) için hedef takibi yapılmış ve Şekil 8'de görüldüğü gibi, optimum örnekleme alan değerinde, $n=4$, hedefin tespit edildiği görülmüştür.



Şekil 6: Optimum $D(x)=0.02$ değerinde hedef takibi



Şekil 7: Optimum $R=10$ değerlerinde hedef takibi



Şekil 8: Optimum örnekleme alan değerinde, $n=4$, hedef takibi

Kilit nokta tanımlayıcıları için gerekli her örnek noktanın temsil edileceği yön sayısı, r 'nin ($r=2-10$ arası), hedef takibine etkisi de araştırılmıştır. Şekil 9'da görüldüğü gibi, optimum değer olan 8'de, hedefin tespit edildiği görülmüştür. Yön sayısı arttığında veya azaldığında ise istenilen hedef takibinde sapmalar tespit edilmiştir. Yön sayısının değişmesi, tanımlanan büyüklüğü, toplam kilit nokta sayısını değiştirmemekle birlikte hedef takibinin doğruluğunu etkilemektedir.

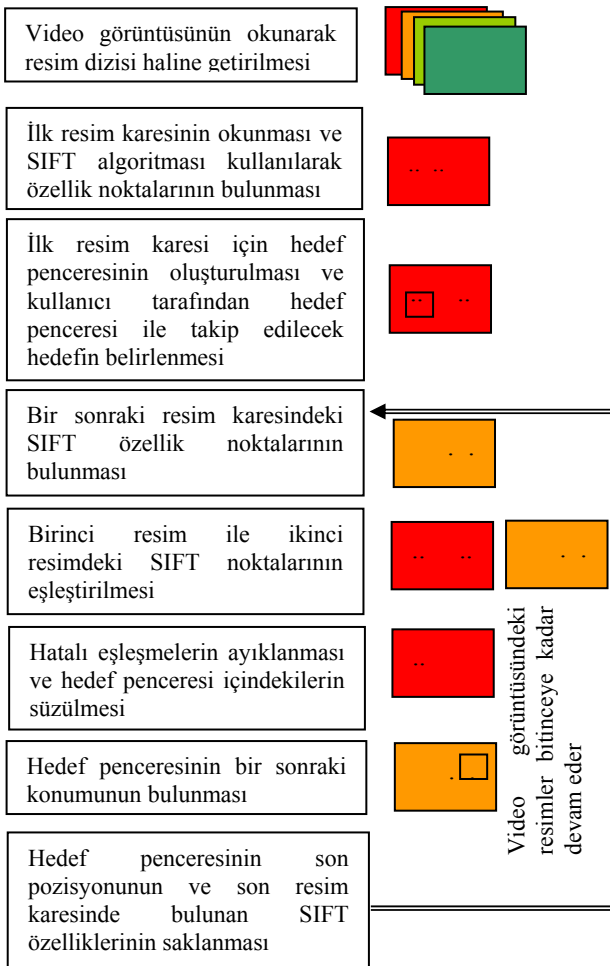
Bulunan sonuçlar, bize Lowe'un optimum değeri olarak atadığı değerlerin doğru olduğunu göstermektedir.



Şekil 9: Optimum $r=8$ değerinde hedef takibi

4. Hedef Takibinde Kullanılan Programın Akış Şeması

Yukarıda, sabit/hareketli kamera ile, sabit/değişken arka plandaki, sabit/hareketli hedefin takibinde kullanılan programın iş akış şeması Şekil 10'da verilmiştir.



Şekil 10: Hedef Takibinde Kullanılan Programın Akış Şeması

5. Sonuçlar

Hedef sabit, kamera hareketli, döngüsel değişimin olmadığı ortamda; hedef sabit, kamera hareketli, döngüsel değişimin kameranın eksenini etrafında olduğu ortamda; hedef ve kamera hareketli, döngüsel değişimin olmadığı ortamda; hedef sabit, kamera hareketli, döngüsel değişimin nesnenin etrafında olduğu ortamda; helikopterden çekilen video görüntüsünde rasgele seçilen hedef ile, değişken arka plan ortamında hareketli hedef ile hedef takibi başarı ile gerçekleştirilmiştir.

Yapılan testler sonucunda, SIFT parametrelerine ait optimum değerler tespit edilmiştir. Bu değerlerin literatürdeki değerler ile uyumlu olduğu gözlenmiştir. Optimum SIFT parametreleri; $s=3$, $D(x)=0.02$, $R=10$, $n_x n_y=4 \times 4$, $r=8$ 'dir.

Ayrıca hedef takibinde oluşabilecek hatalı eşleşmeleri yok etmek için programda yeni bir düzenleme yapılmış, bu düzenleme ile hedef takibinin doğruluğu önemli ölçüde artırılmıştır. Yeni program ile, yukarıdaki optimum SIFT parametre değerlerine yakın değerlerde de hedefe ulaşma olasılığının arttığı gözlenmiştir.

Resim çözünürlüğünün artması ile, SIFT kilit nokta sayısının, SIFT kilit noktalarının tespit süresinin ve hedefi bulma süresinin exponansiyal olarak arttığı gözlenmiştir. Ayrıca resim çözünürlüğünün artması ile, hedef takibinin doğruluğu ve hassasiyeti de artmıştır. Bilgisayar işlemci hızı artırıldığında ise SIFT kilit noktalarını bulma süresinin ve hedefi bulma süresinin azaldığı tespit edilmiştir.

Sonuç olarak, SIFT algoritması kullanılarak, sadece alınan görüntüler üzerinden hedefin ve kameranın hareketlerine karşı dayanıklı noktalar elde edilmiştir. SIFT algoritması ek algoritmalar ile desteklendiğinde daha dayanıklı sonuçlar alınabileceği görülmektedir. İşlemci tabanlı (CPU) platformlar yerine grafik tabanlı (GPU) platformların paralel işlem yapabilen, FPGA tabanlı gömülü sistemler ile birlikte kullanılması; SIFT algoritmasına çok daha fazla hız ve etkinlik kazandıracaktır. Hareketli (uçak, füze, araç,vb.) cisimlerden alınan görüntüler kullanılarak oluşturulan programın, gömülü hale getirilmesi ile hareketli cisimlerin istenilen hedefi takip etmesi sağlanabilecektir.

6. Kaynaklar

- [1] Özgen N., "Bilgisayar Kontrollü Hedef Takibi", **Gazi Üniversitesi, FBE**(Danışman: M. Sarıtaş), Ankara, 2008
- [2] Shi, J. And Malik, J., "Normalized cuts and image segmentation", *IEEE Trans. Patt. Analy. Mach. Intell.*, 8, 888–905 (2000).
- [3] Moravec, H., "Visual mapping by a robot rover", *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 598–600 (1979).
- [4] Harris, C. And Stephens, M., "A combined corner and edge detector", *4th Alvey Vision Conference*, 147–151 (1988).
- [5] Kanade, T., Collins, R., Lipton, A., Burt, P., And Wixson, L., "Advances in cooperative multi-sensor video surveillance", *Darpa IU Workshop*, 3–24 (1998).
- [6] Lowe, D., "Distinctive image features from scale-invariant keypoints", *Int. J. Comput. Vision*, 2, 91–110 (2004).