

AKÜ FEMÜBİD 18 (2018) 015901 (1118-1129)

AKU J. Sci. Eng. 18 (2018) 015901 (1118-1129)

DOI: 10.5578/fmbd.67513

Araştırma Makalesi / Research Article

İnsansız Kara Araçlarından Kamera ile Görüntülenen Hareketli Nesnelerin Sınıflandırılması Amacıyla Geliştirilen Görüntü İşleme Tabanlı Yöntemlerin Karşılaştırılması

Barış Gökçe¹, Güray Sonugür²¹Necmettin Erbakan Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Mekatronik Mühendisliği Bölümü, KONYA²Afyon Kocatepe Üniversitesi, Teknoloji Fakültesi, Mekatronik Mühendisliği Bölümü, AFYONKARAHİSAR

bgokce@konya.edu.tr, gsonugur@aku.edu.tr,

Geliş Tarihi: 23.01.2018 ; Kabul Tarihi:10.10.2018

Öz

Bu çalışmada, hareketli platform olarak kullanılan insansız kara araçlarının seyir halindeyken rotaları üzerindeki bir kısım hareketli nesnelerin tespiti için görüntü işleme tabanlı olarak geliştirilen iki yardımcı sistemin performans karşılaştırması yapılmıştır. Yardımcı sistemlerin geliştirilmesinde küresel konumlandırma sisteminden sağlanan coğrafi konum verileri kullanılmıştır. Çalışmada coğrafi konumlarına göre modelleri çıkartılan arka plan görüntüleri referans model ve anlık model olarak iki farklı şekilde incelenmiştir. Referans modeller varsayılan arka plan (hareketli nesne içermeyen arka plan) görüntülerinden elde edilirken, anlık modeller ise gerçek zamanlı arka plan görüntülerinden elde edilmiştir. Karşılaştırma işlemi sonucu belirlenen silüetler yapay sinir ağları kullanılarak sınıflandırılmıştır. Sınıflandırma için 5 giriş ve 3 çıkışlı bir yapay sinir ağı (YSA) mimarisi uygulanmıştır. Mimarinin son katmanında softmax fonksiyonu kullanılmıştır. Yapay sinir ağında her bir çıkış yaya, otomobil ve bisikletli olarak 3 farklı nesne sınıfını temsil etmektedir. İlk yöntemde iki boyutlu görüntü çakıştırma yöntemi kullanılarak referans ve anlık görüntü çerçeveleri hizalanmıştır. Çerçeve farkı yöntemi ile aynı hizaya getirilen çerçevelerin aritmetik farkları alınarak görüntüye giren dinamik nesnelerin silüetleri elde edilmiştir. İkinci yöntemde arka plan görüntüleri bloklara bölünmüştür. Model oluşturmak için ortalama parlaklık değerleri kullanılmıştır. Referans ve anlık modellerin karşılaştırma işlemi bloklar arasındaki Öklid uzaklıkları hesaplanarak gerçekleştirilmiştir. Yapılan deneysel çalışmalarda; uzaklığa bağlı olarak %94 ile %71 arası doğru sınıflandırma oranları elde edilmiştir.

Anahtar kelimeler

İnsansız kara Araçları;
Görüntü İşleme; Nesne
Tespit, Nesne
Sınıflandırma

Performance Comparison of Image Processing Based Methods Developed to Classify Moving Objects Captured by Camera from Unmanned Land Vehicles

Abstract

In this study, performance comparison of two assistant systems were developed on the basis of image processing for detecting moving objects along the route of unmanned ground vehicles used as mobile platform. The geographical position data provided from global positioning system was used to develop on the assistance systems. In the study, modeled background images according to their geographical locations are handled in two different ways as reference model and instant model. Reference models are derived from default background (background without moving objects), and instant models are derived from real-time background images. The results of the comparison are classified using artificial neural network techniques. In the first method, the reference and instant frames are aligned using two-dimensional image registration methods. By subtracting aligned frames with the frame differencing method, the silhouettes of the moving objects in the scene are obtained. In the second method, the background images are divided into blocks. The average gray-level values are used to create the model. The comparison of the reference and instant models are performed by calculating the Euclidean distances between the blocks. For the classification of the silhouettes 5-input and 2-output artificial

Keywords

Unmanned ground
vehicles ;
Image processing;
Image detection, Image
classification

neural network (ANN) architecture is applied. In this artificial neural network, each output represents a different class of objects as pedestrians, cars and bikes. In experimental studies; the farthest distances determined with a success rate of 85% were determined.

© Afyon Kocatepe Üniversitesi

1. Giriş

Hareketli platform olarak öngörülen insansız kara araçları (İKA) başta uzay çalışmaları (Snorrason et al. 1999), askeri hizmetler (Albekord et al.), kişisel yardım ve destek hizmetleri (Kosuge et al. 2000), toksik temizlik işlemleri (Velasco-Arjona and Leque De Castro) ve tur rehberliği (Philippsen and Siegwart 2003) gibi uygulamalarda kullanılmaktadır. (Gopalakrishnan et al. 2005). Yukarıda bahsedilen işlemlerin İKA'lar tarafından gerçekleştirilebilmesi için nesne tespit etme ve nesne konumunu belirleme gibi bir kısım özelliklere sahip olmaları gerekmektedir. Lidar, ladar ve sonar tipi algılayıcılar, kızılötesi ve ses sinyali bazlı algılayıcılar ile kamera sistemlerini içeren algılayıcıların kombinasyonlarını bir arada barındıran algılayıcılar bu amaçla kullanılabilir. Bu çalışmada algılayıcı olarak kameralar kullanılmıştır. Kameralardan alınan verilerin en iyi şekilde değerlendirilmesi için bir kısım tekniklere ihtiyaç duyulmaktadır. Bu tekniklerden birisi Dijital Görüntü İşlemdir. Son zamanlarda otonom İKA'ların yaygın bir şekilde kullanılmaya başlanmıştır. Bu nedenle kameralardan elde edilen gerçek zamanlı görüntü çerçeveleri kullanılarak nesne tespiti, nesne konumunun belirlenmesi ve nesne tanıma çalışmaları da buna paralel olarak artmıştır. (Shimizu and Yamamoto 2006, Gamez and Devy 2013). Bu kapsamda hava gözetleme sistemleri (Yu and Medioni 2007), araç sürüş yardımı sistemleri (Kong et al. 2010), insansız deniz altı araçları (Foresti and Gentili 2000) vb. konularda dijital görüntü işleme çalışmaları yoğunlaşmıştır.

Dijital görüntü işleme teknikleri ile nesne tanıma işlemlerinde öncelikle nesnelerin doğru bir şekilde tespit edilmesi gerekir. Nesne sıralı çerçeveler arasında sabit ya da hareketli olabilir. Sıralı çerçeveler arasında sabit olarak duran nesneler için pek çok çalışma yapılmıştır ancak hareketli nesnelerin tespiti için sıralı görüntü çerçeveleri arasında gizlenmiş olan bilgi akışını tespit edebilmek gerekir. Bu bilgi akışı; görüntü çerçeveleri arasındaki piksellerin gri seviye ve RGB renk geçişleri belirlenerek, piksellerin oluşturduğu

sınır çizgileri, öbeklenmeler ve bölütlenmiş alanların oluşturduğu öznitelikler belirlenerek tespit edilebilir. (Chen 2008).

Bir İKA için algılayıcı kombinasyonlarından aldığı verileri yorumlaması büyük önem taşır. Bu verilerin hacim olarak en büyüğü ve en değerlisi kameralardan aldığı görüntülerdir. İKA'lar bu görüntü çerçevelerinden elde ettikleri verileri kullanarak bir sonraki hamlelerine karar verirler. Bu sebeple kameralar nesne tespit ve tanıma uygulamaları bilgisayar görme sistemlerinin en önemli parçalarından biridir.

Hareketli platformlarda bulunan kameralardan elde edilen görüntülerden nesne tespiti ile sabit kameralardan elde edilen görüntülerden nesne tespiti arasında büyük farklar bulunur. Çünkü sabit kameralardan alınan ardışık görüntü çerçevelerinde arka plan sürekli aynı kalır. Hareketli kameralardan elde edilen görüntülerdeki arka plan ise hemen hemen her çerçevede değişir ve bir sonraki çerçeve için referans olarak kullanılamaz. Bu nedenle kullanılacak olan yöntemlerin görüntü arka planındaki değişimleri dengeleyebilmesi gerekir.

Literatürde sabit kameralar ile nesne tespit ve tanıma işlemleri için geliştirilen pek çok çalışma bulunmaktadır. Bunlardan bazıları; arka plan tahmin yöntemleri (Jarraya et al. 2010, Spagnolo et al. 2006), optik akış tabanlı yöntemler (Yu et al. 2012, Kong 2010) ve sıralı çerçevelerin farklarının alınması bazlı yöntemlerdir (Weng et al. 2010). Bu çalışmalarda kameraların sabit olması nedeniyle genelde kararlı sonuçlar alınmaktadır. Ancak görüntü işleme tekniklerinde genel sorunlar olan titreme, gölgelerin olumsuz etkisi, ışık değişimleri ve ağaçların sallanma hareketleri gibi görüntülerin azaltılması gerekir.

Hareketli kameralar vasıtasıyla yakalanan ardışık görüntü çerçevelerinde kamera lensinin hareketi nedeniyle arka plan sürekli olarak değiştiği için ~~Bu nedenle~~ sıralı görüntü çerçeveleri arasında konumsal ve kamera açısına göre değişen açısal farklar bulunur. Doğru tespit için öncelikle sıralı iki çerçeve arasındaki bu türlü açısal ve konumsal değişimlerin dengelenmesi gerekir. Bu işlemlerde kullanılan yöntemlere görüntü çakıştırma (image

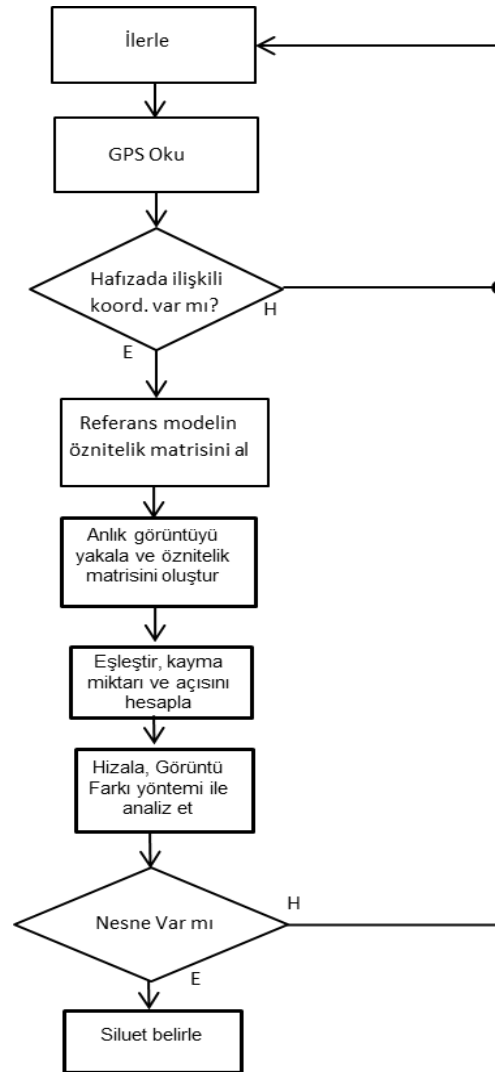
boyutunda skaler değerlerden oluşan bir vektör olarak ifade edilmektedir. Bu çalışmada SURF (Bay et al. 2006), Harris (Harris and Stephens 1988) ve FAST (Rosten and Drummond 2006) öznelik çıkartma yöntemleri kullanılmıştır. Çalışmanın ilerleyen bölümlerinde arka plan görüntüsünün karakteristiğine göre bu yöntemlerin performansları karşılaştırılmıştır. Referans görüntüler belirli yatay eksen aralıklarında ve farklı görüş açılarda alınmıştır. Referans görüntülerdeki özneliklerin görüş açısına ve yatay eksenindeki konumuna göre değişmeyen en güçlü olanları seçilerek diğer öznelikler kaldırılmıştır. Böylece her bir K_i koordinatından alınan referans görüntü kendine özgü benzersiz bir matris ile temsil edilebilmiştir. Bu matris referans model olarak adlandırılmış ve elde edildiği GPS koordinatı ile ilişkilendirilerek veri tabanına kaydedilmiştir. Bu şekilde bilgisayar sistemlerinde işlenmesi gereken görüntü matrisi sayısı ve görüntü matris boyutları küçültülmüştür. Böylece bu alanda yapılan çalışmalarda oldukça kısıtlayıcı bir etken olan gerçek zamanlı işlemlerde harcanan bilgisayar kaynaklarında tasarruf sağlanmış ve çalışma hızı arttırılmıştır.

2.1.2. Referans model ile Anlık Modelin Karşılaştırılması

Bu yöntemde gerçekleştirilen işlem adımları Şekil 2’de verilmiştir. Karşılaştırma işlemlerinin başlangıcında araç belirli zaman aralıklarında GPS koordinatlarını okuyarak belirlenen rotada ilerler. Okuduğu coğrafi koordinatı veri tabanında bilgileri mevcut olan koordinatlar ile karşılaştırır. Bu koordinat daha önceden herhangi bir referans model ile ilişkili olarak kaydedilen koordinatlardan biriye (K_i) bu konuma ait referans model matrisi ile anlık görüntünün model matrisi eşleştirilir. İki boyutlu çerçeve hizalama yöntemi (Gonzalez et al. 2009) kullanılarak referans model ile anlık model arasındaki kayma vektörü hesaplanır. Kayma vektörü kullanılarak anlık görüntü ile referans görüntü birbirine hizalanır ve sonrasında çerçeve çıkarma yöntemi uygulanır. Sonuçta elde edilen görüntüdeki piksel öbekleri analiz edilerek anlık görüntüde bir nesne bulunup bulunmadığı tespit edilir ve döngüye baştan başlanır.

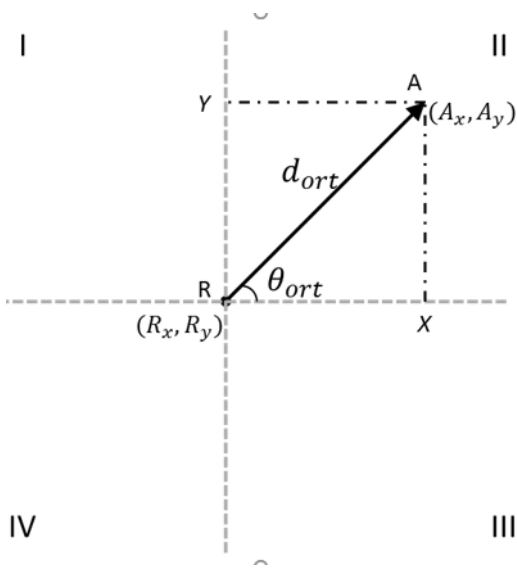
Bu çalışmada referans ve anlık görüntülerin öznelikleri en iyi performansı sağlayan SURF

yöntemiyle çıkarılmıştır. Anlık alınan coğrafi koordinat bilgisi ile hafızadaki koordinat bilgisi uyuştuğu anda İKA’daki kamera ile arka planın görüntüsü yakalanır. Anlık görüntünün SURF öznelik çıkartma algoritması ile kendine has anlık öznelik matrisi oluşturulur. Bu anlık model olarak tanımlanan matristir. Böylece hem model matris hem de anlık matris satır vektörler haline getirilmiştir. Bu adımdan sonra referans model matris ile anlık matrisin değerleri eşleştirilerek aralarındaki iki boyutlu kayma vektörü bulunur. İki matrisin eşleştirilmesi amacıyla model matrisin her bir elemanı için anlık matris ile eşleşen en yakın mesafede bulunan elemanı belirlenir. Vektörler arası doğrusal uzaklık belirleme için Öklid uzaklığı (Euclidean Distance - ED) yöntemi kullanılmış ve model matrisin her bir elemanına en yakın anlık matris elemanlarından oluşan bir $Uzaklık_{i,j}$ matrisi oluşturulmuştur.



Şekil 2. GPS destekli çerçeve hizalama ve çerçeve farkı yöntemi akış çizelgesi.

Bu işlem sonucunda elde edilen kayma vektörlerinin görsel ifadeleri Şekil 3'de verilmiştir. Her bir vektör bir genlik ve bir kayma açısından oluşmuştur. Referans ve anlık görüntü arasındaki toplam kayma miktarını bulmak üzere tüm genlik ve açı değerlerinin ortalaması alınmaktadır. Bu ortalama değerler referans ve anlık görüntülerin birbirine tekrar hizalanması için kullanılacaktır. Eğer kayma vektörü doğru olarak belirlenemez ise görüntüler doğru olarak hizalanamayacak ve sonucunda elde edilecek olan fark görüntüden doğru bir tespit yapmak mümkün olmayacaktır.



Şekil 3. Kayma vektörü

Şekilde;

R: Referans görüntüye ait piksel,

A: Anlık görüntüye ait piksel,

d_{ort} : Ortalama kayma miktarı,

θ_{ort} : Ortalama kayma açısını,

R_x, R_y : Referans görüntü pikselinin x ve y konumlarını,

A_x, A_y : Anlık görüntü pikselinin x ve y konumlarını ifade etmektedir.

Bu çalışmada koordinat sistemi I, II, III, IV olmak üzere 4 bölgeye ayrılmıştır. Şekilde belirtilen RA vektörü II. bölgede bulunmaktadır. Kayma açısının tespiti amacıyla kayma vektörünün x ve y eksenlerine olan izdüşümlerinin hesaplanmıştır. Bu izdüşüm değerlerinin hesaplanması Denklem 1,2'de verilmiştir.

$$X = d_{ort} \cdot \cos(\theta_{ort}) \quad (2.1)$$

$$Y = d_{ort} \cdot \sin(\theta_{ort}) \quad (2.2)$$

Kayma vektörünü bulmak için yön bilgisini de bulmak gerekmektedir. Referans görüntü piksellerinin koordinat sisteminin merkezi olarak varsayıldığında anlık görüntü piksellerinin yeni koordinatlarının hesabı bulunduğu bölgeye göre değişecektir. Koordinat sistemini bölgelerine göre piksellerin yeni x ve y koordinatları Denklem 2.3 ile Denklem 2.10 arasında ifade edildiği şekilde hesaplanmıştır.

I. bölge:

$$A_x(\text{yeni}): A_x + X \quad (2.3)$$

$$A_y(\text{yeni}): A_y + Y \quad (2.4)$$

II. bölge:

$$A_x(\text{yeni}): A_x - X \quad (2.5)$$

$$A_y(\text{yeni}): A_y + Y \quad (2.6)$$

III. bölge:

$$A_x(\text{yeni}): A_x - X \quad (2.7)$$

$$A_y(\text{yeni}): A_y - Y \quad (2.8)$$

IV. bölge:

$$A_x(\text{yeni}): A_x + X \quad (2.9)$$

$$A_y(\text{yeni}): A_y - Y \quad (2.10)$$

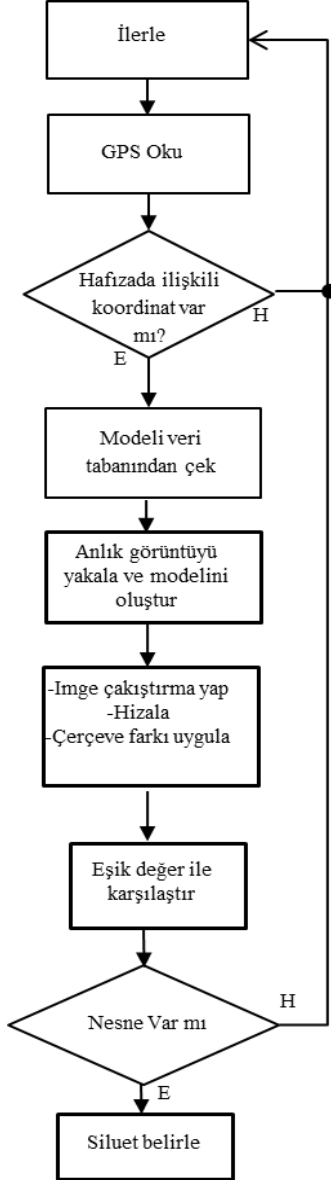
Referans görüntü ile anlık görüntü kayma vektörüne göre hizalandıktan sonra sıralı çerçeve farkı yöntemi uygulanarak iki görüntü arasındaki fark piksel öbekleri ortaya çıkartılır. Bunlar anlık görüntüdeki nesne silüetleridir. Silüetlerin daha net olarak belirlenebilmesi için eşikleme ve morfolojik işlemler uygulanmıştır. İkili görüntülerin elde edilebilmesi için Otsu (Otsu 1979) yöntemi ile çeşitli morfolojik yöntemler uygulanmıştır.

2.2 Gps Destekli Özelleştirilmiş İmge Çakıştırma Yöntemi (Yöntem-2)

Bu yöntemde; GPS destekli çerçeve hizalama ve çerçeve farkı yönteminde yapılan çalışmaya benzer olarak görüntü çakıştırma metodu kullanılmıştır. İmge çakıştırma işlemi referans görüntüler ile anlık görüntüler arasında yapılmaktadır. Ancak görüntü çakıştırma için kullanılan teknik farklıdır. Diğer yöntemlerde olduğu gibi bu yöntem de iki aşamadan oluşmuştur ve GPS yardımı kullanılmıştır. Yöntemde kullanılan algoritmanın akış çizelgesi Şekil 4'de verilmiştir.

2.2.1 Referans Modelin Oluşturulması

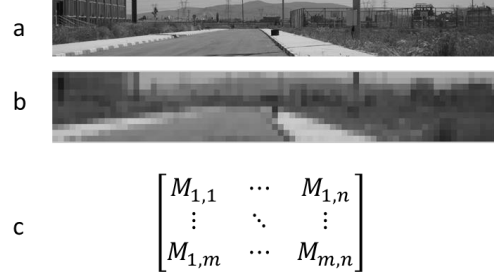
Birinci yöntemdekine benzer olarak Üniversite kampüsünde belirlenen her bir K_i koordinatında yabancı nesne içermeyen arka plan görüntüleri fotoğraflanmıştır ve bunlar referans görüntüler olarak adlandırılmışlardır. Bu çalışmada diğer yöntemden farklı olarak her bir koordinat için sadece 1 adet referans görüntü kullanılmıştır.



Şekil 4. Özelleştirilmiş görüntü çakıştırma yöntemi akış çizelgesi

Öncelikle her bir koordinattan elde edilen referans görüntüler gri seviyeye dönüştürüldükten sonra boyutları parametrik olarak değiştirilebilen dikdörtgen bloklara bölünerek mozaik görüntüler elde edilmiştir. Bu mozaik blokların gri seviye değerlerinin ortalaması o bloğu temsil eden gri seviye değeri olarak atanmıştır. Bu şekilde piksel geçişlerindeki küçük gürültülerin azaltılması ve keskin ton değişimlerinin görüntü üzerindeki

etkisinin azaltılmıştır. Bu şekilde her koordinat için varsayılan arka planı temsil eden tek bir referans görüntü R_i elde edilmiştir. Bu referans görüntünün matrisi görüntünün fotoğraflandığı GPS koordinatının referans modelidir. Rota boyunca tüm koordinatlardan elde edilen referans modeller koordinatlar ile ilişkilendirilerek veri tabanına kaydedilir. Yapılan işlemlerin görsel ifadesi Şekil 5'de verilmiştir.



Şekil 5. (a) Referans görüntü (b) Mozaik görüntü (c) Referans görüntüye ait matris.

Referans ve anlık modeller arasında karşılaştırma yapılarak kayma vektörünün hesaplanabilmesi için ortak özniteliklerin çıkarılması gerekmektedir. Bu yöntemde bir önceki yöntemde kullanılan SURF, Harris ve FAST vb. öznitelik çıkartma yöntemleri yerine mozaik görüntüdeki bloklar arasındaki gri seviye değişimlerini modelleyecek 4 adet yeni öznitelik oluşturulmuştur. Bunlar 2 boyutlu bir görüntü x : satır ve y : sütun olmak üzere $f(x,y)$ fonksiyonu olarak ifade edildiğinde;

1. Her bir bloğun yatay ve dikey eksen boyunca gri seviye değişimini gösteren gradyan genliği (Denklem 3.1),
2. Gradyan açısı değeri (Denklem 3.2),
3. Her bir bloğun birinci türev değeri,
4. Her bir bloğun ikinci türev değeridir.

İki boyutlu fonksiyonların gradyanları x , y noktalarındaki en yüksek değişimin gerçekleştiği yönü işaret ederler ve bir vektör olarak ifade edilirler. Bu vektör Denklem 3.1'de ifade edilmiştir. Bu denklemde g_x ve $(\partial f/\partial x)$ ifadeleri $f(x,y)$ görüntü fonksiyonunun x değişkenine göre kısmi türevini, g_y ve $(\partial f/\partial y)$ ifadeleri ise $f(x,y)$ görüntü fonksiyonunu y değişkenine göre kısmi türevini göstermektedir. Gradyan vektörü 2 adet değerden oluşmuştur. Bu değerler; değişimin büyüklüğünü ifade eden genlik ve değişimin yönünü ifade eden açı değerleridir. Genlik bilgisi Denklem 3.2'de, açı bilgisi ise Denklem 3.3'de belirtilmiştir.

$$\text{gradyan}(f) = \nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad 3.1$$

$$\text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad 3.2$$

$$\theta(x, y) = \tan^{-1}[g_y/g_x] \quad 3.3$$

Böylece her bir K_i koordinatına ait referans model için bir öznitelik matrisi oluşturulmuş ve veri tabanına kaydedilmiştir. Öznitelik matrisi her blok için 1x4 boyutundaki vektörlerden oluşmaktadır.

2.2.2. Referans model ile Anlık Modelin Karşılaştırılması

Bu adım gerçek zamanlı hareket sırasında gerçekleşir. İKA hareketine başladığı anda GPS koordinat bilgisini alarak ilerlemeye başlar ve döngüsel olarak bu bilgiyi okur. Aldığı her koordinat bilgisini veri tabanında kayıtlı olanlar ile karşılaştırır ve eşitliği yakaladığı an kameradan anlık görüntüyü alır. Anlık görüntüyü referans görüntüdekiyle aynı işlem adımlarından geçirerek modeller. Böylece anlık görüntüye ait model A_i oluşturulmuş olur. Sonrasında referans ve anlık görüntü arasındaki kaymayı bulabilmek için iki model karşılaştırılır. Aynı K_i coğrafi konumuna ait referans modellerin oluşturduğu dizi R_i olarak sembolize edildiği takdirde karşılaştırma işlemi R_i ile A_i arasındaki Öklid uzaklığının hesaplanmasıyla gerçekleştirilebilecektir. Hesaplama sonucunda referans model ile anlık model arasında blok başına vektörel uzaklıklar çıkartılır. Yani bir bakıma referans model ile anlık model arasında birbirine en çok benzeyen bloklar arası parlaklık seviyesi değişimleri bulunur. En kısa mesafeli bloklar arasındaki uzaklık değerleri ile $U_{x,y}$ matrisi oluşturulur. Bu matris birbirine en çok benzeyen dikdörtgen blokların bulunmasında kullanılacaktır. Bu çalışmada $U_{x,y}$ matrisinde en çok tekrar eden değerler iki model arasındaki kayma miktarı olarak kabul edilmiştir. Kayma miktarı yatay ve dikey eksen kayma değerlerini içermekte olup anlık görüntünün referans görüntüye hizalanması için kullanılmıştır. Aynı coğrafi koordinata sahip referans ve anlık görüntüler hizalandıktan sonra çerçeve farkı yöntemi ile farkları alınarak sahnedeki hareketli nesnenin silueti elde edilmektedir.

2.3 Yapay Sinir Ağları ile Sınıflandırma

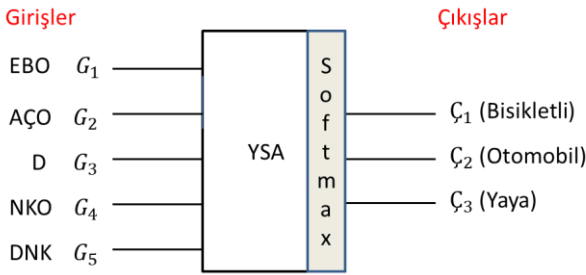
Sınıflandırma işlemi matematiksel bir model ile yapıldığı takdirde veri tabanında kullanılan model ile ilişkili nesne bildirim modellerini tutmak gerekir. Ancak sınıflandırma işleminde YSA türevi yöntemler kullanılacak ise nesne tanıma verileri sinir ağı ağırlıklarında tutulacağı için bir veri tabanına ihtiyaç bulunmamaktadır. Genel olarak nesnelerin doğru olarak tanınması pek çok koşulun bir arada gerçekleşmesiyle yapılabilmektedir. Ancak çoğu zaman ideal şartlar oluşmaz ve bazı eksiklikler ve belirsizlikler meydana gelebilir. Değişken aydınlanma koşulları, nesnenin bir kısmının bir başka cismin arkasında kalması, aşırı hareketli arka plan, esnek yapıdaki nesneler gibi belirsizlikler yapılacak işlemin başarısını olumsuz yönde etkilemektedirler. Bu nedenlerle matematiksel ve yapısal hesaplama yöntemleri çoğu zaman hız ve doğruluk anlamında istenilen sonucu verememektedir. Buna karşılık paralel çalışan ve birbirleriyle bağlantılı işlem elemanlarından oluşan yapay sinir ağı sınıflandırıcıları insan düşünme ve sınıflandırma sistemini taklit edebilen sistemler olup nesne tanıma işlemlerinde geniş olarak kullanılmaktadırlar. Bu teknikte basit işlem elemanlarının kullanılmasına rağmen bu işlem elemanlarının paralel çalışması sonucunda hız ve doğruluk açısından tatmin edici sonuçlar alınabilmektedir.

Tasarlanan YSA mimarilerinde her bir çıkış farklı bir nesne sınıfını temsil eder. Sınıflandırma işlemi için pek çok algoritma geliştirilmiştir. Bu çalışmada Çok Katmanlı Ağ (Multi Layer Perceptron-MLP) mimarisinde Geri Yayımlı Sinir Ağı (Back Propagation Neural Network-BPNN) algoritması uygulanmıştır. Bu çalışmada 5 girişli ve 2 çıkışlı bir YSA mimarisi uygulanmıştır. Mimaride 50 işlem elemanından oluşan bir gizli katman bulunmaktadır. Girdi değişkenleri aşağıda sıralanmıştır.

- I. En/Boy oranı (EBO): Tespit edilen nesnenin piksel sayısı cinsinden eni ile boyu arasındaki orandır. Ölçek değişmezliği için oran olarak alınmıştır. Aksi takdirde nesnenin kamera lensine uzaklığına göre farklı değerler elde edilecektir.
- II. Alan/Çevre oranı (AÇÖ): Tespit edilen nesnenin piksel sayısı olarak toplam alanı ile çevresi arasındaki orandır. Yine ölçek değişmezliği için oran tercih edilmiştir.

- III. Dairesellik (D): 0 ile 1 arası skalar bir değer olup "1" değerine yakın olması daireselliğe yaklaştığını "0" değerine yakın olması ise çizgiselliğe yaklaştığını göstermektedir.
- IV. Nesne/Kutu oranı (NKO): Bu değer nesnenin piksel cinsinden alanının kutunun toplam alanına oranıdır. Oran olarak alındığı için ölçek değişmezliğine sahiptir.
- V. Dip Nesne/Kutu oranı (DNK): Bu değer nesnenin en alt 1/20'lik kısmının nesne/kutu oranıdır. Bu oran insan-otomobil ayırımında belirleyici bir faktördür.

Çıktılar ise **001**: Yaya, **010**: Otomobil, **100**: Bisikletli şeklinde tasarlanmıştır. Tasarlanan YSA mimarisi kabaca Şekil 6'da verilmiştir.



Şekil 6. Tasarlanan YSA mimarisi

Tasarlanan YSA'nın eğitimi için yaklaşık 1100 adet görüntü kullanılmıştır. Bu görüntülerde farklı beden ölçülerinde, farklı pozisyonlarda (ön, yan, arka, yürüyüş pozisyonu, durma pozisyonu vs.) ve farklı kıyafetlerde yayalar, farklı pozisyonlarda ve modellerde araç ve bisikletliler bulunmaktadır. Bu görüntülerden bir kısmı Şekil 7'de verilmiştir.



Şekil 7. YSA'nın eğitiminde kullanılan görüntülerin bir kısmı

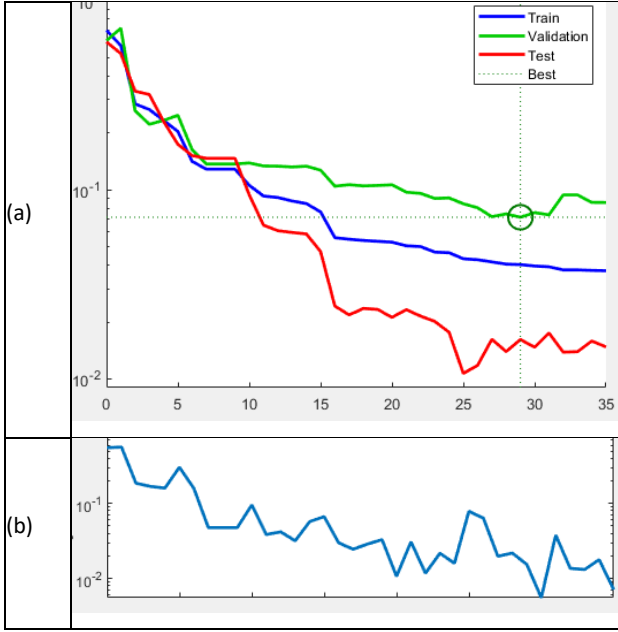
Morfolojik görüntü işleme yöntemleri kullanılarak bu görüntülerin ön plan ve arka plan segmentasyon işlemleri yapılmış ve her nesnenin (yaya, bisikletli, otomobil) silüetleri ikilik görüntü şeklinde elde edilmiştir. Bu ikilik görüntü kullanılarak her nesne için YSA'na girecek olan ve Şekil 6'da verilen girdi parametrelerinin sayısal değerleri elde edilmiştir. Bu sayısal değerlerden her nesne grubundan rastgele seçilen değerler Çizelge 1'de verilmiştir.

Çizelge 1. Girdi parametrelerinin bir kısmına ait sayısal değerler

Girdi nesnesi	Alan/çevre oranı	Nesne piksel/kutu piksel	Nesne
Yaya	25,95	0,62	
Otomobil	31,01	0,81	
Bisikletli	31,311	0,533	

YSA mimarisinde 1 adet 10 nöronlu oluşan gizli katman kullanılmıştır. Eğitim için verilerin %70'i kullanılmış, test ve doğrulama işlemleri için verilerin %30'u eşit miktarlarda kullanılmıştır.

Kameradan 15m uzaklıktaki nesnelere içeren bir veri seti kullanılarak yapılan deneysel YSA çalışmasından elde edilen performans grafiği Şekil 8'de ve gradyan iniş grafiği Şekil 9'da verilmiştir.



Şekil 8. (a) YSA'nın performans grafiği, (b) gradyan iniş grafiği

Bu veri setinde YSA 35 iterasyonda varsayılan hata miktarının altına düşerek eğitimini tamamlamıştır. YSA'da gerçekleştirilen öğrenme işlemi sonuçlarının değerlendirilmesi için YSA daha önce eğitim veri setinde kullanılmamış test veri seti ile teste tabi tutulmuştur. Eğitim ve test işlemleri sonucunda oluşan karmaşıklık matrisi Çizelge 2.'de verilmiştir.

Çizelge 2. Karmaşıklık matrisi

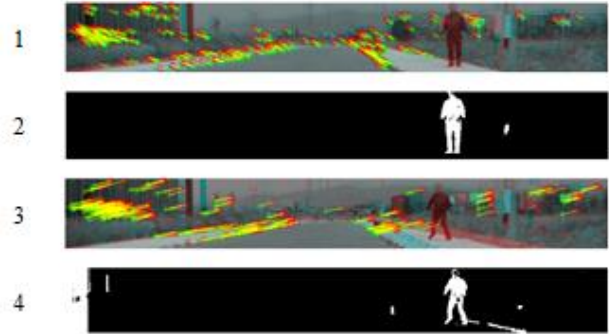
	Bisikletli	Otomobil	Yaya
Bisikletli	93	5	2
Otomobil	5	92	1
Yaya	9	7	88

Karmaşıklık matrisinde sütunlar tahmin edilen sınıfları, satırlar ise gerçek sınıfları göstermektedir. Buna göre; bisikletli değerleri incelenirse 100 adet bisikletliden 93 adeti doğru olarak sınıflandırılmış, 5 adet bisikletli figürü otomobil olarak sınıflandırılmış ve 2 tanesi de yaya olarak sınıflandırılmıştır. En kötü sonuçların alındığı yaya sınıfında ise toplam 105 adet figürden 88 tanesi doğru sınıflandırılmış, 16 adeti ise bisikletli ve otomobil olarak sınıflandırılmıştır. Bu veri setinde Bisikletli için doğru sınıflandırma oranı %93, otomobil için %94 ve yaya için ise %84 olarak gerçekleşmiştir.

4. Sonuç

Nesne tespit ve tanıma çalışmalarında tespit edilecek nesnelerin yapıları büyük önem taşımaktadır. Rijit yapıda olan nesnelere daha kolay tespit edilebilirken esnek yapıda olan nesnelerin tespitinde zorluklar ile karşılaşmaktadır. Bu nedenle deneysel çalışmalarda; rijit yapıları tespit etmek amacıyla otomobiller, esnek yapıları tespit etmek amacıyla da insanlar kullanılmıştır. Bisikletliler hem rijit yapıda hem de esnek yapıda unsurlar barındırdığından her iki kategoriye de girmektedir. Yöntemleri karşılaştırabilmek için her iki yöntemde de bir kısım deneysel çalışmalar yapılmıştır.

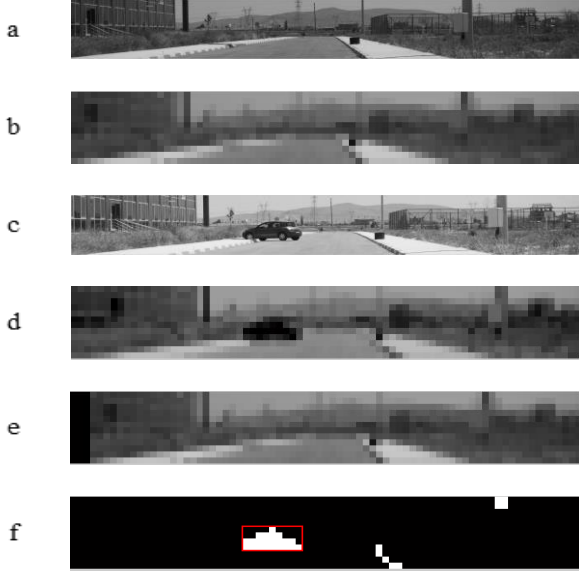
GPS destekli çerçeve hizalama ve çerçeve farkı yöntemi (Yöntem-1) ile yapılan bir deneysel çalışmada herhangi bir K_i koordinatındaki kameraya 15m mesafede bulunan yaya figürleri incelenmiştir. Şekil 9(1)'de referans görüntü ile anlık görüntü arasındaki kayma verilmiştir. Şekil 9(2)'de sıralı çerçeve çıkarma yöntemi uygulanarak ve çeşitli morfolojik işlemlerden geçirilerek elde edilen ikilik görüntü verilmiştir. Bu örnek referans ve anlık görüntü çiftinde yatay ve dikey kaymanın yüksek olmadığı görülmektedir. Şekil 9 (3-4)'te ise yatay ve dikey kaymanın daha yüksek olduğu referans ve anlık görüntüler verilmiştir. Kayma miktarı yüksek olduğu için Şekil 9(4)'te verilen ikilik görüntüde silüetin etrafında bulunan gereksiz piksel öbeklerinin tam olarak yok edilemediği görülmektedir.



Şekil 9. (1) K_i koordinatında düşük kayma miktarına sahip anlık resim, (2) elde edilen nesne silüeti, (3) yüksek kayma miktarına sahip anlık resim, (4) elde edilen nesne silüeti

GPS destekli özelleştirilmiş görüntü çakıştırma yöntemi (Yöntem-2) ile yapılan deneysel çalışmada herhangi bir K_i koordinatına 22m uzaklıktaki bir otomobil kullanılmıştır. Her bir işlem basamağında elde edilen görüntüler Şekil 10'da verilmiştir. Yapılan deneysel çalışmalardan anlaşıldığı üzere insan bedenine göre daha rijit ve hacimce büyük

olan otomobillerin insana göre daha uzaktan tespit edilebildiği, kayma miktarı azaldıkça nesne silüetlerinin daha net belirlenebildiği, kayma miktarının yüksek değerlerde çıkmaması için herhangi bir koordinattaki referans model çeşitliliğinin artırılması gerekmektedir



Şekil 10. İşlem basamaklarında elde edilen görüntü örnekleri

Şekil 10'un (a) bölümünde koordinatın referans görüntüsü, (b) bölümünde referans görüntünün bloklara ayrılmış hali, (c) bölümünde anlık görüntü, (d) bölümünde anlık görüntünün bloklara ayrılmış hali, (e) bölümü kayma miktarı hesaplandıktan sonra anlık görüntünün referans görüntüye hizalanmış hali, (f) bölümünde ise referans görüntü ile anlık görüntü arasındaki farkın ikilik görüntüye dönüştürülmüş durumu verilmiştir.

Yöntemlerin performanslarının değerlendirilmesi için yaya, otomobil ve bisikletlilerin 10m, 20m ve 25m mesafelerdeki başarılı sınıflandırma verileri hesaplanmıştır. Aynı ölçme tiplerinin kullanıldığı Yöntem-1 ve Yöntem-2'de 3264x4928 çözünürlükte elde edilen sonuçların karşılaştırmalı listesi Çizelge 3'de verilmiştir.

Çizelge 3. 3264x4928 çözünürlükte 1. ve 2. yöntemin performans karşılaştırması

Yöntemler	Nesne Tipi	Başarı Oranları (%)		
		15m	20m	25m
Yöntem-1	Yaya	84	81	76
Yöntem-2	Yaya	82	81	71
Yöntem-1	Otomobil	94	85	83
Yöntem-2	Otomobil	94	83	82
Yöntem-1	Bisikletli	93	83	80
Yöntem-2	Bisikletli	91	82	79

Görüntü çözünürlüklerinin düşürülmesi kullanılan kameranın maliyetini düşürdüğü için önemlidir ancak yüksek çözünürlükte elde edilen sonuçların yanında kabul edilebilir seviyelerde sonuç elde edilmesi gerekir. Ayrıca görüntü piksel sayılarındaki azalma nedeniyle hesaplama maliyetleri de düşecektir. Düşük çözünürlükte (480x640) yapılan çalışmalarda gözlemlenen sonuçlar ise Çizelge 4 'de verilmiştir.

Çizelge 4. 480x640 çözünürlükte 1. ve 2. yöntemin performans karşılaştırması

Yöntemler	Nesne Tipi	Başarı Oranları (%)		
		15m	20m	25m
Yöntem-1	Yaya	82	80	73
Yöntem-2	Yaya	82	76	70
Yöntem-1	Otomobil	94	84	82
Yöntem-2	Otomobil	93	84	81
Yöntem-1	Bisikletli	93	82	80
Yöntem-2	Bisikletli	90	81	79

Alınan sonuçlara göre; nesnelerin kameraya olan uzaklıkları arttıkça başarılı nesne sınıflandırma oranı düşmektedir. Nesne kameradan uzaklaştıkça ikilik görüntülerde gürültü artmakta ve silüeti net olarak elde edilememektedir. Sınıflandırma başarısı otomobil gibi rijit yapıda olan nesnelere yüksek iken yaya gibi esnek yapıda olan ve değişik formlara girebilen nesnelere ise daha düşüktür. YSA'nın çok farklı formlarda yayaları öğrenmek zorunda kalması sınıflandırma başarısını düşürmüştür. Ayrıca kameraya uzak bir mesafeden yayalar topluluk halinde yürüdüklerinde YSA tarafından bisikletli ya da otomobil olarak değerlendirilebilmektedir. Sınıflandırma başarısı olarak incelendiğinde bisikletliler, yaya ve otomobiller arasında bir yere konumlanmıştır. Çünkü bisikletliler hem rijit hem de esnek unsurlar içermektedirler. Yöntemlerin doğru sınıflandırma oranları incelendiğinde Gps Destekli Çerçeve Hizalama ve Çerçeve Çıkartma Yönteminin (Yöntem-1) Gps Destekli Özelleştirilmiş İmge Çakıştırma Yöntemine (Yöntem-2) göre daha başarılı olduğu görülmektedir. Yöntem-1'de görüntü bir bütün olarak ele alınıp standart öznelikleri çıkartılmaktadır. Bu işlem başarıyı arttırmaktadır ancak Yöntem-1 iki adımdan oluşmaktadır. Önce çerçeve hizalama ve sonrasında çerçeve çıkartma işlemleri yapılmaktadır. Bu nedenle daha başarılı sonuçlar alınmasına rağmen Yöntem-2'ye göre hesapsal maliyeti yüksektir. Yöntem-2'de ise görüntü önce karesel bloklara ayrılıp sonrasında öznelik çıkartma işlemi yapıldığı için bir kısım görüntü özellikleri kaybolabilmektedir.

Bu da sınıflandırma işleminde başarıyı azaltan bir unsurdur.

Güvenlik sistemlerinde kameralardan alınan görüntülerin anlık olarak izlenmesi insan faktörüne dayalı dolayısıyla da insan hatasına açık olan bir işlemdir. Bu çalışma ile kameralardan alınan görüntülerin izlenmesine ve kaydedilmesine dayalı güvenlik sistemlerine farklı bir platform kazandırılmıştır. Bu platform kameraları sabit ve pasif olmaktan çıkartıp hareketli hale getirmiştir. Kameralardan uzak olduğu için güvenlik sistemleri tarafından net olarak izlenemeyen konumlara kameraların İKA'ları aracılığıyla yaklaştırılabilmesi sağlanmıştır. Bu şekilde güvenlik tedbirleri sağlanabileceği gibi uzak konumlara yeni kameralara olan gereksinim ortadan kaldırılmıştır. Ayrıca sistem üzerinde bulunan yapay zeka unsurları ile güvenlik tedbirleri anlık olarak tespit edilebilmekte, istendiği takdirde geri bildirimde bulunulabilmekte bu şekilde de insan faktörüne bağımlılık azaltılmaktadır. Tüm bu özellikleri ile birlikte geliştirilen sistem İKA'ları üzerine yerleştirildiğinde insansız bir güvenlik devriyesi olarak kullanılabilir.

Teşekkür

Bu çalışma Afyon Kocatepe Üniversitesi Bilimsel araştırma projeleri komisyonunun 14.FEN.BİL.37 nolu projesi kapsamında yapılmış çalışmadır. Destek sağlayan ilgili kuruluşlara teşekkür ederiz

5.Kaynaklar

- Albekord, K., Watkins, A., Wiens, G., Fitz-Coy, N., and Lin, K. C. ,2004. Multiple-agent surveillance mission with non-stationary obstacles. *Proceedings of 2004 Florida Conference on Recent Advances in Robotics*, 1-5.
- Bay, H., Tuytelaars, T. and Van Gool, L., 2006. Surf: Speeded up robust features. *Computer vision–ECCV*, 404-417.
- Chen, X., 2008. Application of matlab in moving object detecting algorithm. *Future BioMedical Information Engineering, 2008. FBIE'08. International Seminar*, 114-117.
- Cheraghi, S.A.,2012. "Moving Object Detection Using Image Registration for a Moving Camera Platform," 23–25
- Foresti, G. L. and Gentili, S.,2000. A vision based system for object detection in underwater images. *International Journal of Pattern*

Recognition and Artificial Intelligence, **14(02)**, 167-188.

- Gonzalez, R.C., Woods, R.E. and Eddins, S.L. 2009. *Digital Image Processing using Matlab*, Gatesmark, USA
- Gopalakrishnan, A., Greene, S. and Sekmen, A., 2005. Vision-based mobile robot learning and navigation. *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop* ,48-53.
- Harris, C. and Stephens, M. 1988. A combined corner and edge detector. *Alvey vision conference* , **15(50)**, 10-5244.
- Jarraya, S. K., Hammami, M. and Ben-Abdallah, H., 2010. Accurate background modeling for moving object detection in a dynamic scene. *Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference*, 52-57.
- Kim, J., Ye, G. and Kim, D., 2010,. Moving object detection under free-moving camera. *Image Processing (ICIP), 2010 17th IEEE International Conference*,4669-4672.
- Kong, H., Audibert, J. Y. and Ponce, J. (2010). Detecting abandoned objects with a moving camera. *IEEE Transactions on Image Processing*, **19(8)**, 2201-2210.
- Kosuge, K., Sato, M., and Kazamura, N., 2000. Mobile robot helper. *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference*, 1,583-588.
- Marquez-Gamez, D. A. and Devy, M., 2013. Active vision-based moving objects detection from a Motion Grid. *Mobile Robots (ECMR), 2013 European Conference*, 373-378.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, **9(1)**, 62-66
- Philippsen, R. and Siegwart, R., 2003. Smooth and efficient obstacle avoidance for a tour guide robot. In None (No. LSA-CONF-2003-018).
- Rosten, E. and Drummond, T. 2006. Machine learning for high-speed corner detection. *Computer Vision–ECCV*, 430-443.
- Sappa, A. D., Dornaika, F., Gerónimo, D. and López, A., 2008. Registration-based moving object detection from a moving camera. *Proc. on Workshop on Perception, Planning and Navigation for Intelligent Vehicles, Nice, France*
- Shimizu, S., Yamamoto, K., Wang, C., Satoh, Y.,

- Tanahashi, H. and Niwa, Y., 2006. Moving object detection by mobile Stereo Omni-directional System (SOS) using spherical depth image. *Pattern analysis and applications*, **9(2-3)**, 113-126.
- Snorrason, M., Norris, J., & Backes, P. ,1999. Vision based obstacle detection and path planning for planetary rovers. Proceedings of SPIE 3693, 13th Annual AeroSense conference, Orlando, FL
- Spagnolo, P., Leo, M. and Distanto, A., 2006. Moving object segmentation by background subtraction and temporal analysis. *Image and Vision Computing*, **24(5)**, 411-423.
- Velasco-Arjona, A., and de Castro, M. L., 1997. Fully robotic method for characterization of toxic residues. *Analyst*, **122(2)**, 123-128.
- Yu, Q. and Medioni, G., 2007. Map-enhanced detection and tracking from a moving platform with local and global data association. *Motion and Video Computing, 2007. WMVC'07. IEEE Workshop*, 3-3.
- Yu, X., Chen, X. and Gao, M.,2012. Motion detection in dynamic scenes based on fuzzy c-means clustering. *Communication Systems and Network Technologies (CSNT), 2012 International Conference*, 306-310.
- Weng, M., Huang, G. and Da, X., 2010. A new interframe difference algorithm for moving target detection. *Image and Signal Processing (CISP), 2010 3rd International Congress* **1**, 285-289.