

HAAR-BENZERİ ÖZNETELİKLER VE BASAMAKLI HIZLANDIRMA KULLANARAK TÜMYÖNLÜ KAMERALAR İLE BİNEK ARAÇ TESPİTİ

CAR DETECTION WITH OMNIDIRECTIONAL CAMERAS USING HAAR-LIKE FEATURES AND CASCADED BOOSTING

Hakkı Can Karaimer, Yalın Baştanlar

*Bilgisayar Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü
{cankaraimer,yalinbastanlar}@iyte.edu.tr*

ÖZETÇE

Bu bildiri tümyönlü görüntüler üzerinde binek araç tespiti için bir yaklaşım sunmaktadır. İlk olarak, normal kamera görüntüleri için geleneksel Haar-benzeri öznitelikler ve basamaklı hızlandırma yöntemi açıklanmıştır. Daha sonra, bu yöntemi tümyönlü görüntüler üzerinde uygulamak için, tümyönlü imgelerden panoramik imgeler oluşturulmuştur. Bu şekilde araç tespiti tek imge üzerinde, tümyönlü görüntüden çok sayıda perspektif imge oluşturmaya gerek kalmadan gerçekleştirilmiştir. Ayrıca iki farklı panorama oluşturma şekli tartışılmış ve küresel kesitli olanların silindirik panoramalardan daha uygun olduğu görülmüştür. Gerçek tümyönlü imgeler ile yapılan araç tespit deneylerimiz sunulmuştur.

ABSTRACT

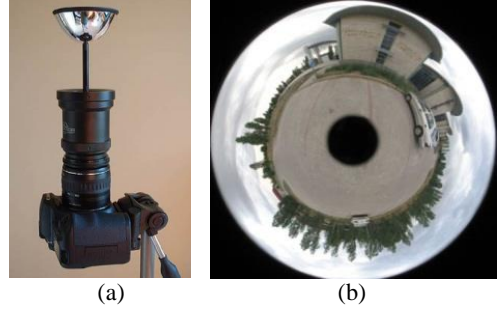
This paper presents an approach to detects cars in omnidirectional images. We first go through the conventional method of using Haar-like features and cascaded boosting for conventional camera images. Then, to apply this method for omnidirectional cameras, we generate panoramic images from omnidirectional ones. In this way we perform car detection on a single image without generating numerous perspective images from the omnidirectional view. We also discuss two different ways of panoramic image generation and conclude that spherical profile panoramas are more convenient than cylindrical panoramas. We present our car detection experiments on real omnidirectional images.

1. GİRİŞ

Tek bir imgede 360° görüntüleme sağlayan tümyönlü kameralar, özellikle son on yılda, güvenlik [1] ve 3B geriçatım [2] da dahil olmak üzere çeşitli alanlarında yaygın olarak kullanılmaya başlanmıştır. Tümyönlü imge elde etmek için en sık kullanılan yöntem kamera önüne dışbükey aynalı aparatlar yerleştirmektir (Şekil 1). Bu görüntüleme sistemleri için katadioptrik (catadioptric) sıfatı kullanılır ve tek görüş noktalı (single-viewpoint) sistemlerin imge oluşum geometrisi daha kolay modellenebilir [3]. Aynasız geniş açılı görüntü alımı için ise balıkgözü lensler kullanılmakta ve yarımküre görüş açısı sağlanabilmektedir.

Geniş görüş açısı sayesinde az sayıda tümyönlü kamera daha çok sayıda standart kameranın yaptığı işi yapabilir. Fakat tümyönlü kameralarla nesne tespiti konusunda, özellikle de yaya ve araç tespiti gibi trafik güvenliğine hizmet edecek konularda henüz kayda değer miktarda çalışma yapılmamıştır.

Bu bildirideki çalışmalar TÜBİTAK tarafından 113E107 nolu proje kapsamında desteklenmiştir.



Şekil 1: (a) Katadioptrik tümyönlü kamera elde etmek için bir ayna aparatı geleneksel bir kamera önüne yerleştirilir. (b) Katadioptrik kamera ile elde edilmiş bir örnek imge.

Tümyönlü kameralar ile nesne tanıma konusunda bir çalışmada [4], bir robota belirli nesnelerin resimleri verilmiş, ve gezindiği ortamda bu nesnelere bulması istenmiştir. Bu çalışmada tam olarak aynı nesne arandığından, doğrudan SIFT [5] ile nokta eşleme yapılmış, araba, insan v.b. gibi bir nesne sınıfı tanımlanmamıştır. [6]'da yazarlar katadioptrik tümyönlü imgelerde yüz bulma için, tümyönlü imgeyi panoramik imgeye çevirmişler ve doğrudan Viola-Jones yüz bulma metodunu [7] uygulamışlardır. Tümyönlü imgeleri panoramik imgelere çevirmeden doğrudan Haar-benzeri öznitelikler çıkaran ve yüz bulmada kullanan bir çalışma [8]'de verilmiştir. Fakat uygulanan yöntem baktığı pencereler için %5 yanlış-ortı sonuç döndürmektedir. Bu oran, Viola ve Jones tarafından önerilen, kayan pencereler kullanarak imgede yüz tespiti yapan bir metodun sahip olması gereken $1 \cdot 10^{-6}$ yanlış-ortı oranından çok yüksektir. Dolayısıyla, doğrudan tümyönlü imgede uygulanma avantajına rağmen kullanılması anlamlı değildir.

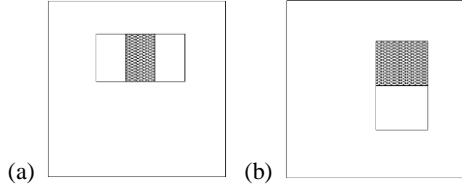
Viola ve Jones [7] Haar-benzeri öznitelikleri ve basamaklı hızlandırmayı beraber kullanarak yüz bulma üzerine kilometre taşı denebilecek bir çalışma yapmışlardır. Bu yaklaşım, özniteliklerin seçiminde farklı bir yöntem izlenerek ve araba bulmaya yönelik olarak [9]'da kullanılmıştır. Bu çalışmada yazarlar UIUC'nin [10] binek araçların yandan görünüşlerini içeren eğitim ve test setlerini kullanmışlardır.

Bu bildirinin ikinci bölümünde, Viola ve Jones'un önerdiği ve pek çok çalışmanın belkemiğini oluşturan Haar-benzeri öznitelikler ve basamaklı hızlandırma kullanan yöntem tarif edildikten sonra üçüncü bölümde bizim tümyönlü imgelerde bu yöntemi nasıl kullandığımız açıklanmıştır. Dördüncü bölümde ise deney sonuçları aktarılmıştır.

2. HAAR-BENZERİ ÖZNETELİKLER VE BASAMAKLI HIZLANDIRMA YÖNTEMİ

2.1. Haar-Benzeri Öznitelikler

Haar-benzeri öznitelikler imgedeki dikdörtgenel bölgelerin piksel değerlerinin toplamının farkı olarak ifade edilir. *İki-dikdörtgen öznitelik*, iki dikdörtgenel bölgenin farkını, *üç-dikdörtgen öznitelik* ise dışarıda kalan iki dikdörtgenden ortadaki dikdörtgenin çıkarılmasını ifade eder (Şekil 2). Daha farklı, örneğin dört-dikdörtgen öznitelikler de kullanılabilir.



Şekil 2: Haar-benzeri öznitelikler. Beyaz dikdörtgenel alanların piksel toplamından gri dikdörtgenel alanların piksel toplamı çıkarılır. (a) Üç-dikdörtgen öznitelik. (b) İki-dikdörtgen öznitelik. [7]

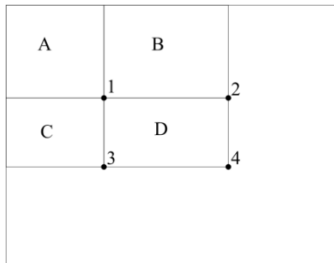
Biz çalışmamızda sadece Şekil 2'de verilen öznitelikleri kullandık ve 80x32 piksel büyüklüğündeki modelimizi yandan görünen binek araçları tespit edecek şekilde eğittik. Belirtilen öznitelik tipleri için 80x32'lik pencereye sığan, değişik en-boy oranına sahip tüm varyasyonlarının farklı pozisyonlarda bulundurulması ile elde edilen toplam öznitelik sayısı 73,000 dolayındadır. Sayıyı makul tutmak amacıyla, dikdörtgen bölgeler imge üzerinde bir yerine iki piksel atılarak hesaplama yapılmıştır.

2.2. İntegral İmge

Dikdörtgen özniteliklerin hızlı hesaplanabilmesi için integral imge kullanılır. Integral imgenin (x,y) konumundaki değeri, (x,y) 'nin üstünde ve solunda kalan alanın piksel değerleri toplamıdır. Şu formülle ifade edilir:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

Burada $ii(x,y)$ integral imgeyi, $i(x,y)$ ise orijinal imgeyi göstermektedir. İntegral imge kullanılarak herhangi bir dikdörtgenel toplam dört referansla hesaplanabilir (Şekil 3). Benzer şekilde, Şekil 2b'deki öznitelik, iki komşu dikdörtgenel toplam içerdiği için altı referansla, Şekil 2a'daki fark ise sekiz referansla hesaplanabilir.



Şekil 3: A dikdörtgeni içindeki piksel değerleri toplamı, integral imgenin 1 konumundaki değeridir. 2 konumundaki değer A+B, 3 konumundaki değer A+C, 4 konumundaki değer A+B+C+D'dir. D içindeki piksel değerleri toplamı 4+1-(2+3) olarak hesaplanır [7].

2.3. Sınıflandırıcının Eğitilmesi

AdaBoost, ilk önerildiği haliyle [11] bir sınıflandırıcıyı hızlandırmak için kullanılmış, fakat Viola ve Jones [7] tarafından önerilen ve bizim de benimsediğimiz halinde, sınıflandırıcıyı eğitmenin yanısıra özniteliklerin seçilmesi için de kullanılmıştır.

Nesneyi imgenin farklı ölçeklerle küçülmüş hallerinde kayan pencereler şeklinde aradığımızdan (Şekil 4) sınıflandırıcı da bir pencereyi araba veya değil şeklinde ayırt etmek üzerine eğitilir. Her bir pencereden gelen 73,000 dikdörtgen özniteliklerin çok küçük bir kümesi etkili bir sınıflandırıcı için kullanılır. Bu amaç doğrultusunda, AdaBoost algoritmasındaki zayıf sınıflandırıcılar bir öznitelikmiş gibi düşünülerek pozitif ve negatif örnekleri en iyi şekilde ayıran öznitelikler aranır. Her zayıf sınıflandırıcı, minimum sayıda yanlış sınıflandırılmış örnek veren eşik değerli bir fonksiyondur. Bir zayıf sınıflandırıcı $h(x,f,p,\theta)$; bir öznitelik f , bir eşik değeri θ ve eşitsizlik işaretinin yönünü belirten bir polarite p 'den oluşur. Burada x imgenin 80x32'lik bir penceresidir.

$$h(x, f, p, \theta) = \begin{cases} 1 & p \cdot f(x) < p \cdot \theta \text{ ise} \\ 0 & \text{aksi halde} \end{cases} \quad (4)$$

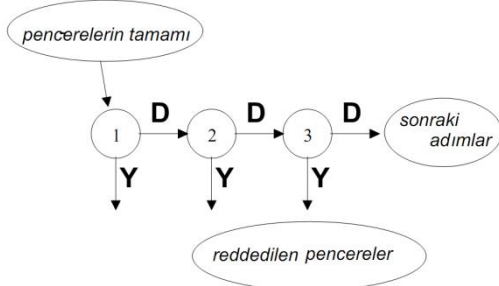
Esas sınıflandırıcı ise bir dizi zayıf sınıflandırıcının birleşmiş halidir. Her bir zayıf sınıflandırıcı (öznitelik) eklendiğinde, eğitim kümesinde doğru sınıflandırılan örneklerin ağırlıkları azaltılır, yanlış sınıflandırılan örneklerinki ise artırılır. Güncellenmiş ağırlıklarla en az hatayı veren yeni bir zayıf sınıflandırıcı seçilir ki bu önceki sınıflandırıcıların yanlış sonuç ürettiği örneklerde daha az hata yapan farklı bir özniteliktir. Bu şekilde farklı hatalar için özelleşmiş zayıf sınıflandırıcılar birleştirilerek nihai bir sınıflandırıcı oluşur. Hedeflenen doğru tespit ve yanlış alarm oranlarına ulaşıldığında öznitelik ekleme durdurulur.

2.4. Basamaklı Sınıflandırıcı

Basamaklı hızlandırmadaki temel amaç küçük ve dolayısıyla daha etkili, neredeyse tüm pozitif örnekleri tespit ederken pek çok negatif pencereyi reddeden hızlı sınıflandırıcılar oluşturabilmektir. Yanlış-artı pencerelerin çoğunluğu ilk aşamada az öznitelik içeren nispeten basit sınıflandırıcılar ile reddedilebilir. Daha düşük yanlış-artı oranı elde etmekle görevli karmaşık sınıflandırıcılar ise sadece ilk aşamayı geçen pencereler için çalıştırılır. Böylece, toplam süre oldukça azalır [7]. Tarif edilen basamaklı sınıflandırıcı yapısı Şekil 5'te gösterilmiştir. İlk sınıflandırıcıdan gelen bir pozitif sonuç, yine yüksek tespit oranı için ayarlanmış ikinci bir sınıflandırıcıyı tetikler. İkinci sınıflandırıcı ilk aşamada elenemeyen yanlış-artıları elemek üzerine eğitilir. İkinci sınıflandırıcıdaki bir pozitif değer ise üçüncü sınıflandırıcıyı tetikler ve bu böyle devam eder. Birkaç aşama işlemden sonra yanlış alarm pencerelerinin sayısı büyük ölçüde azalmıştır. Herhangi bir noktadaki negatif bir çıktı o pencerenin derhal reddine sebep olur. Bu basamaklardaki sınıflandırıcılar AdaBoost ile eğitilir.



Şekil 4: Kayan pencereler şeklinde arama



Şekil 5: Tüm pencerelere uygulanan basamaklı sınıflandırıcının şematik gösterimi. Nispeten basit ilk sınıflandırıcı az bir işlem yüküyle yanlış-artı örnekleri büyük ölçüde eler. Sonraki katmanlar ek hesaplama ile kalan yanlış-artıları eler. [7]

3. YÖNTEMİN TÜMYÖNLÜ GÖRÜNTÜLERDE UYGULANMASI

Normal kamera imgeleri ile eğitilmiş bir nesne modeli kullanılırsa, integral imgeyi ve Haar-benzeri öznelikleri tümyönlü imgeler üzerinde hesaplamak (başta trigonometrik fonksiyonlar olmak üzere) ağır bir hesap yükü getirmektedir. Bu ek yük tümyönlü imgenin panoramik imgeye çevriminden daha ağırdır. Ayrıca, dikdörtgen alanların oranları tümyönlü imgede sabit olmadığından Şekil 2'deki öznelilikler doğrudan kullanılamaz. Bu nedenle çalışmamızda, tümyönlü imge panoramik imgeye çevrilip nesne tespit işlemi orada gerçekleştirilmiştir.

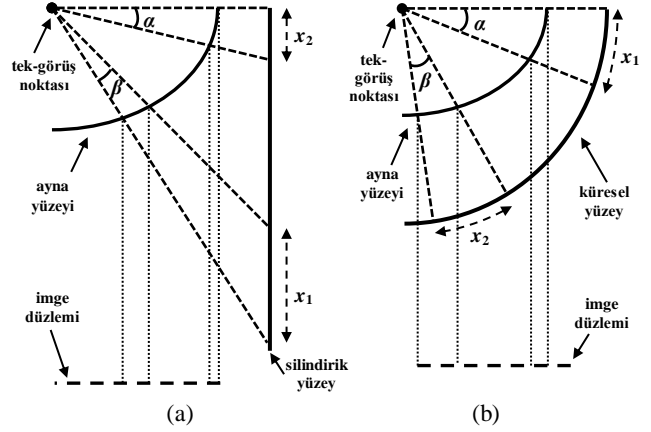
Nesne modelini doğrudan tümyönlü imgelerle eğitmek bir alternatif gibi görünmesine rağmen, kayan pencerenin şekli nesnenin tümyönlü imgedeki boyutuna ve merkeze yakınlığına göre değiştiğinden eğitilecek sabit bir model tanımlamak pratikte mümkün değildir. Şekil 6'da gösterildiği üzere kameraya yakın bir araba çok bükülmüş iken uzaktaki bir araba daha düzdür, yani standart kamera imgesine daha yakındır. Aradaki fark sadece büyüklük olmadığından, bir model tanımlayıp ölçükle büyütüp/küçültmek çözüm değildir.



Şekil 6: Tümyönlü imgede kameraya farklı uzaklıkta bulunan ve siyah çerçeve ile gösterilen iki araçtan, büyük olan daha bükülmüş bir pencere ile aranmalı, küçük olan ise daha düz bir pencere ile aranmalıdır.

Panoramik imgenin oluşturulacağı yüzey için bazı alternatifler mevcuttur. Sık kullanılan silindirik yüzey, nesne tespiti için en uygun yüzey değildir. Göz (kamera odağı) seviyesinden

aşağıya inildikçe nesnelere uzar ve normal kamera imgeleri ile eğitilen modele benzememeye başlar. Şekil 7a'da kesit çiziminde gösterildiği gibi silindir üzerinde aşağı hareket ettikçe eşit açılar için daha uzun mesafeler harcanır ($\alpha=\beta$, $x_1>x_2$). Daha uygun bir yüzey olan küre kesitinde ise (Şekli 7b) eşit açılar eşit mesafeler ile temsil edilir ($\alpha=\beta$, $x_1=x_2$). Normal kamera görüntülerine bu açıdan daha yakın olan küresel kesitli yüzey nesne tespiti için de daha uygundur. Bununla ilgili deney sonuçları bir sonraki bölümde verilmiştir. Silindirik ve küresel kesitli örnek panoramik imgeler Şekil 8'de verilmiştir.



Şekil 7: Silindirik ve küresel yüzeyde panoramik imge oluşumu



Şekil 8: Örnek silindirik yüzey (üstte) ve küresel kesitli yüzey (altta) ile oluşturulan panoramik imgeler

4. DENEYLER

Önceki bölümde önerdiğimiz küresel kesit ile hazırlanmış panoramik imgelerin silindirik panoramik imgelere göre nesne tespitinde daha başarılı sonuçlar verdiğini doğrulamak üzere bu iki yöntemi karşılaştıran deneyler gerçekleştirildi.

UIUC veriseti[10] ile Haar-benzeri öznelilikler ve basamaklı hızlandırma kullanarak modelimizi oluşturduk. Bu verisetindeki negatif örnek sayısı(500) kayan pencereler şeklinde çalışan bir metodu eğitmek için yeterli değildir, çünkü $1 \cdot 10^{-3}$ ve aşağısında yanlış-artı oranlarına ulaşabilmek için daha geniş bir negatif kümeye ihtiyaç vardır. Bu amaçla her bir basamak sonucunda eğitilen sınıflandırıcı araç içermeyen imgeler içerisinde gezdirilip yanlış-artıları toplamak suretiyle daha geniş bir negatif küme elde edilmiştir. Tüm eğitim basamaklarının toplamında 550 pozitif örneğin yanında 8400 negatif örnek kullanılmıştır. Yeni basamaklar, bir önceki basamakta izin verilen yanlış artıların %90'ını eleyecek şekilde oluşturulmuştur. Bu şekilde eğitilen modelimizde sırasıyla 8, 19 ve 31 zayıf sınıflandırıcı içeren üç basamak yer almıştır.

Aynı tümyönlü imgeden elde edilmiş küresel kesitli ve silindirik panoramik imgeler üzerinde araç tespiti yapılmıştır. Arabaların pozisyonları etiketlenmiş pencere ile algoritmanın pozitif sonuç üreten bir penceresinin örtüşme oranı (7)'de verilen formülle hesaplanıp, %50'nin üzerindeyse doğru-artı (da) olarak kaydedilir. Örtüşmesi yeterli olmayan pozitif sonuçlar yanlış-artı (ya), algoritma tarafından bulunamayan etiketler ise da yanlış-eksi (ye) olarak sayılır.

$$\ddot{O} = \frac{\text{Alan}(\text{tespit penceresi} \cap \text{etiketlenen pencere})}{\text{Alan}(\text{tespit penceresi} \cup \text{etiketlenen pencere})} \quad (7)$$

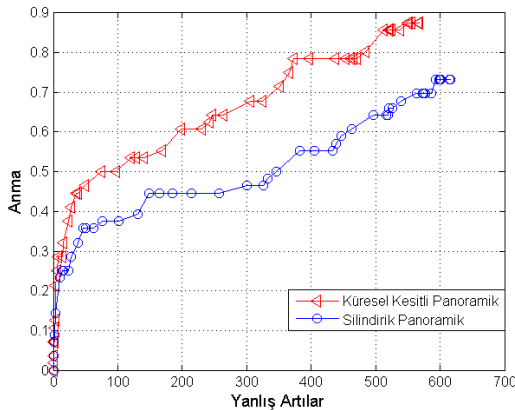
Son sınıflandırıcıdan geçen ve birbirleriyle örtüşen pencereler için de benzer bir yol izlenmiş, %50'nin üzerinde örtüşmelerde sadece yerel maksimum alınmıştır.

25 farklı imgenin ters yansıtılmış (ayna simetrisi) halleri ile toplamda 50 imge ile yaptığımız deneyin sonuçları, yanlış-artı sayısına karşı anma (da/(da+ye)) eğrisi olarak Şekil 8'de verilmiştir. Grafikteki noktalar farklı eşik değerlerine denk gelmektedir. Küresel kesitli imgelerdeki başarımın daha iyi olduğu görülmektedir. Aradaki fark, küresel kesitle oluşturulmuş imgelerde araçların eşik değerini daha kolay geçebilmesinden ötürüdür. Görsel örnek vermek amacıyla, Şekil 10 ve 11'de eşik değeri, sırasıyla en yüksek skorlu iki ve üç pencere geçecek şekilde belirlendiğinde bulunan pencereler görülmektedir. Silindirik panoramik imgede araç bu pencereler arasında değildir.

Basamaklı sınıflandırma işleminin ilk aşamasını oluşturan sekiz zayıf sınıflandırıcı ile tüm pencerelerin ortalamada %96.3'ü elenmiştir. Dolayısıyla imgeden çıkan pencerelerin yüzde 4'ünden daha azı için ilerleyen basamaklar çalışmıştır. Bu da bir imge için işlem süresini 2.17 saniyeden 0.71 saniyeye düşürmüştür.

5. DEĞERLENDİRME

Tümyönlü kameralar ile binek araç tespiti amaçlı bu çalışmamızda, yöntem olarak Haar-benzeri öznelilikler ve basamaklı hızlandırma kullanılmış, nesne tespiti ise tümyönlü imgelerden elde edilen panoramik imgelerde kayan pencerelerin eğitilen modele yakınlığı ölçülerek yapılmıştır. Basamaklı hızlandırmanın tek bir sınıflandırıcıya göre kazandırdığı süre gözlemlenmiş ayrıca küresel kesitli ve silindirik panoramik imgeler, tespit başarı oranları kullanılarak birbirleriyle kıyaslanmıştır. Öngörüldüğü üzere küresel kesitli panoramalarda başarı daha yüksek çıkmıştır.



Şekil 9: Silindirik ve küresel kesitli panoramik imgelerin basamaklı sınıflandırıcı deney sonuçları. +∞, -∞ aralığındaki eşik değerlerinde; x eksen, tüm imgeler için toplam yanlış-artı sayısını, y eksen anma değerlerini vermektedir.



Şekil 10: Eşik değeri, en yüksek skorlu iki pencere geçecek şekilde ayarlandığında; küresel kesitli (üstte) ve silindirik (altta) panoramik imgelerin test sonuçları (koyu renkli çerçeve) ve etiketlenen araç (beyaz çerçeve).



Şekil 11: Eşik değeri, en yüksek skorlu üç pencere geçecek şekilde ayarlandığında; küresel kesitli (üstte) ve silindirik (altta) panoramik imgelerin test sonuçları (koyu renkli çerçeve) ve etiketlenen araç (beyaz çerçeve).

6. KAYNAKÇA

- [1] G. Scotti, L. Marcenaro, C. Coelho, F. Selvaggi, C. Regazzoni, "Dual camera intelligent sensor for high definition 360 degrees surveillance", *IEE Proc. Vision Image Signal Processing*, 152(2), 2005.
- [2] M. Lhuillier, "Toward flexible 3d modeling using a catadioptric camera", *Proc. of Computer Vision and Pattern Recognition (CVPR)*, 2007.
- [3] S. Baker, S.K. Nayar, "A Theory of Single-Viewpoint Catadioptric Image Formation", *Int. Journal of Computer Vision*, 35(2), 175-196, 1999.
- [4] M.L. Wang, H.Y. Lin, "Object Recognition from Omnidirectional Visual Sensing for Mobile Robot Applications", *IEEE International Conference on Systems, Man and Cybernetics*, 2009.
- [5] D. Lowe, "Distinctive image features from scale invariant keypoints", *International Journal of Computer Vision*, 60, 91-110, 2004.
- [6] A. Iraqui, Y. Dupuis, R. Bouteau, J. Ertaud, X. Savatier, "Fusion of omnidirectional and PTZ cameras for face detection and tracking", *International Conference on Emerging Security Technologies*, 2010.
- [7] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", *Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [8] Y. Dupuis, X. Savatier, J.Y. Ertaud, P. Vasseur, "A direct approach for face detection on omnidirectional images", *IEEE International Symposium on Robotic and Sensors Environments (ROSE)*, 2011.
- [9] F. Moutarde, B. Stanculescu, A. Breheret, "Real-time visual detection of vehicles and pedestrians with new efficient adaBoost features", *2nd Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV)*, 2008.
- [10] S. Agarwal, D. Roth, "Learning a sparse representation for object detection", *Proc. of the Seventh European Conference on Computer Vision, Part IV*, 113-127, 2002.
- [11] Y. Freund, R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Computational Learning Theory: Eurocolt '95 Springer-Verlag*, 23-37, 1995.