

HOG Temelli Bir Yöntem ile Ölçek ve Yönden Bağımsız Gerçek Zamanlı Nesne Tanıma

Murat Peker¹, Halis Altun², Fuat Karakaya³

^{1,2,3} Elektrik Elektronik Mühendisliği, Niğde Üniversitesi, Niğde

¹ e-posta: murat.pkr@gmail.com ² eposta: halisaltun@nigde.edu.tr

³ eposta: fkarakaya@nigde.edu.tr

Özetçe

Nesne ve örüntü tanıma uygulamalarında, yüksek başarımla elde edilen Histogram of Oriented Gradient (HOG) algoritması yakın bir zamanda literatürde önerilmiştir. HOG algoritması tarafından oluşturulan, basit ancak etkin tanımlayıcılar, ele alınan probleminin çözümünde kullanılan sınıflandırıcıların yüksek başarımla nesne tanıma ve sınıflandırma yapabilmesine imkan sağlamaktadır. Bu çalışmada HOG algoritmasına dayalı olarak endüstriyel bir ortamda geometrik şekillerin tanımlanmasını mümkün kılacak bir yöntem sunulmuştur. Yöntem ile, HOG tabanlı sınıflandırma işleminin başarımlarını artırmak için, HOG tanımlayıcılarının elde edilmesi aşamasında, algoritmanın başarımlarını düşüren arka plan gürültüsünün varlığına dikkat çekmekte ve bu gürültünün etkisizleştirilmesi için, gömülü sistemlere de uyarlanabilecek şekilde, bir ön işlem önerilmektedir. Elde edilen sonuçlar bu yöntemin kullanılması durumunda sistem genel başarımlarında %200 artış elde edilebileceğini göstermektedir. Ayrıca önerilen yöntem ile endüstriyel uygulamalar için sınıflandırma probleminin çözüm hızının %23 artırılacağı da gösterilmiştir. Önerilen yöntemin gömülü sistemlere uyarlanabilecek bir yapıda olması da yöntemin bir başka avantajı olarak görülebilir.

1. Giriş

Endüstriyel ve askeri uygulamaları olan imge işleme ve örüntü tanıma ile alakalı çalışmalar literatürde yoğun olarak verilmekte ve mevcut algoritma ve sınıflandırıcıların performansını artırıcı yönde çeşitli öneri ve yöntemler araştırmacılar tarafından sunulmaktadır. Son yıllarda Shashua ve diğerleri [1] ve Dalal ve Triggs [2] tarafından önerilen HOG algoritması bu yönde yapılan başarılı çalışmalara örnek olarak verilebilir.

Nesne ve örüntü tanıma için yaygın olarak kullanılmaya başlanan Histogram of Oriented Gradient (HOG), bir çok konuda ve farklı şartlar altında çok yüksek bir başarımla çalışabilmektedir. HOG, ilk olarak Shashua ve diğerleri [1] tarafından yaya tanıma sistemlerinde kullanılabilir tanımlayıcılar olarak önerilmiştir. Dalal ve Triggs [2] bu yeni tanımlayıcıları başarı ile kompleks ortamlarda insan tanıma probleminin çözümüne uyarladılar. Bu basit fakat etkin tanımlayıcı, başarılı uygulamalarından dolayı, literatürde yoğun bir ilgi kazanmış ve birçok uygulamada kullanılmaya başlanmıştır [3][4].

Bizim çalışmamızda ele aldığımız şekil tanıma problemi ise bir çok sınıflandırma ve tanıma uygulamasında çözülmesi gereken temel alt bir problemdir ve El şekli tanıma [5], trafik

işaretleri tanıma [6-9], insan tanıma [6][10], çeşitli objelerin tanımlanması [11] gibi uygulamalarda kullanımı söz konusudur

Çalışmamızda HOG algoritması tarafından oluşturulan tanımlayıcıların kullanılması ile temel geometrik şekillerin ayırt edilmesi gerçekleştirilmiştir. Önerilen yöntem ile açıdan ve boyuttan bağımsız sınıflandırma gerçekleştirmek mümkündür. Böylece literatürde kullanılan farklı algoritmaların başarımlarını artırmak için açı düzeltmek (alignment) ve boyut sabitlemek (scale normalization) gibi gerekli ön işlemlerin uygulanmasına ihtiyaç hissedilmeyecektir. Algoritma ilk olarak arka plan düzensizliklerinin olmadığı ideal şartlar altında tasarlandı. Bu ideal koşullar altında HOG tabanlı algoritma %100 gibi yüksek bir başarımla ulaştığı gösterildi. Ancak ideal şartların söz konusu olmadığı farklı endüstriyel uygulamalar altında önerilen algoritmanın başarımlarının ciddi bir şekilde düştüğü gözlemlendi. Bu düşüşün temel nedeninin arka plan düzensizlikleri (gürültü) olduğu tespit edildi. Bu tespit doğrultusunda HOG tabanlı tanımlayıcılar elde etmeden önce arka plan gürültüsünün giderilmesi için özgün bir algoritma önerildi. Bu algoritma sonucunda endüstriyel uygulamada sistem başarımlarının ciddi bir şekilde artış gösterdiği gözlemlendi. Aynı zamanda bu iyileşme sistem hızında herhangi bir taviz verilmeden gerçekleştirildi ve önerilen yöntem sayesinde sistem hızında ortalama %28'lük bir iyileşme de elde edildi.

Makalenin organizasyonu şu şekilde yapılmıştır; 2. bölümde HOG algoritmasının uygulamamızda kullanılışı, 3. bölümde uygulamamızın ideal şartlar altında geliştirilmesi izah edilmiştir. 4. bölümde endüstriyel ortam altında sistem performansı gözlemlenmiş ve yöntemin iyileştirilmesi için bir arka plan gürültü giderici teknik önerilmiştir. 5. bölümde ise elde edilen sonuçlar değerlendirilmiştir.

2. HOG

Son yıllarda imgedeki piksellerin yönelim (θ) ve büyüklük değerlerinin karakteristiği olarak da adlandırılabilir olan HOG algoritmasının kullanımı bir çok alanda oldukça ilgi görmektedir. HOG kullanımı ilk defa Shashua [1] ve Dalal [2] tarafından önerilmiştir. Bir çok araştırmacı tarafından oldukça ilgi gören HOG yöntemindeki temel amaç imgeyi bir grup lokal histogramlar olarak tanımlamaktır. Bu gruplar, imgenin lokal bir bölgesindeki gradyanların yönelimlerinde, gradyanların büyüklüklerinin toplandığı histogramlardır.

Bir imgenin HOG değerlerinin çıkarılması için gerekli olan hesaplamalar şu şekilde gerçekleştirilmektedir. İlk olarak imgenin, yatay ve dikey Sobel filtreleri uygulanarak, I_x ve I_y olmak üzere kenarları belirlenir (Denklem 1 ve 2). Sobel

filtresi uygulanmış I_x ve I_y imgeleri kullanılarak, gradyan ve bu gradyanların yönelim açılarının hesaplanması için (G ve θ);

$$I_x = I * S_y \quad (1)$$

$$I_y = I * S_d \quad (2)$$

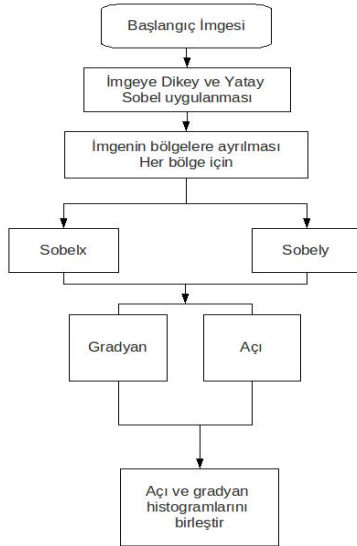
$$|G| = \sqrt{I_x^2 + I_y^2} \quad (3)$$

$$\theta = \arctan \frac{I_x}{I_y} \quad (4)$$

formüllerini kullanılır. Uyguladığımız HOG tanımlayıcısında lokal histogram bölgeleri tanımlanmamıştır. Bunun yerine bütün bir imge tek bir bölge olarak işleme alınmıştır. Uygulamamızda kullanılan HOG algoritmasının işleyiş diyagramı Şekil 1'de görülmektedir.

HOG algoritması uygulanırken oluşturulan histogramda piksellerin yönelim açılarının gruplanması ile daha anlaşılabilir sonuçlar elde edilebilmektedir. Bu gruplama olayı 0-360 aralığında yer alan açısal değerlerin istenilen bir aralığa çekilmesi ile mümkün olmaktadır. Uygulamamızda 12 derecelik 30 aralığa sahip açısal yönelim grupları kullanılmıştır. Çıkarılan HOG tanımlayıcısının değerlendirilmesi kural tabanlı bir sınıflandırma kullanılarak gerçekleştirilmiştir. Kural tabanlı sınıflandırma ile elde edilen kenar sayıları geometrik şeklin tanımlanmasında kullanılmıştır.

Sonraki bölümde arka planı sabit ve homojen bir yapıda olan geometrik şekillerin tanımlanmasına yönelik uygulanan yöntemler ve sonuçlar verilmiştir..

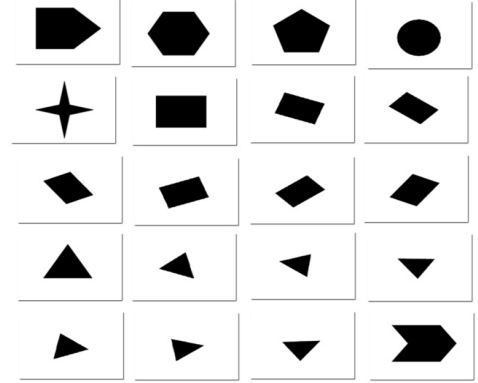


Şekil 1. HOG algoritması akış diyagramı

3. Geometrik Şekillerin Tanınması

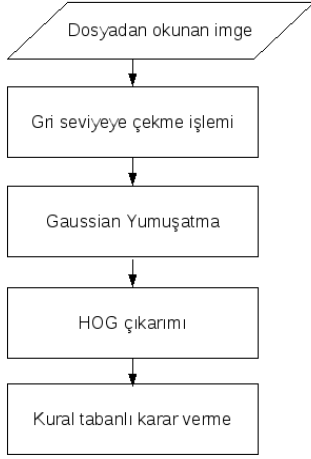
Bu bölümde arka planı sabit renkli ve homojen yapıda olan geometrik şekillerin tanımlanmasını amaçlayan bir yöntem sunulmuştur. HOG algoritmasının hem FPGA gibi gömülü sistemlere uygulanabilirliğinden dolayı hem de imgeye ait yön

bilgisini verebilmesi dolayısı ile de bu uygulamamızda kullanılması uygun görülmüştür. Bu amaçla geometrik şekillerin tanınmasını gerçekleştirecek bir yazılım ilk olarak PC üzerinde gerçekleştirilmiştir. PC üzerinde uygulanan yazılım c++ dili kullanılarak hazırlanmıştır. Uygulamada giriş imgelerini elde etmek için OpenCV kütüphanesinden faydalanılmıştır. Giriş imge kaynağı olarak web cam, imge dosyaları ve video dosyaları kullanılabilir. Arka planı homojen ve sabit renkli test imgelerimiz çizim programları yardımı ile hazırlanmıştır. Test imgelerimiz boyutları değişken geometrik şekiller içermektedir ve 125 adettir. Test imgelerimizden bazı örnekler Şekil 2'de görülmektedir.



Şekil 2. Test imgelerinden örnekler

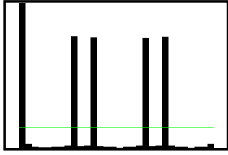
Geometrik şekillerin tanınması için uygulanan sistemde giriş olarak verilen imgeler öncelikle gri seviyeye çekilmektedir. Gri seviyeli imge 7x7'lik bir gaussian filtresi kullanılarak yumuşatılmaktadır. Elde edilen imge üzerinde yatay ve dikey yönlere Sobel filtreleri uygulanarak I_x ve I_y imgeleri hesaplanmaktadır. I_x , I_y imgeleri kullanılarak HOG değerleri hesaplanmaktadır. HOG değerleri 12 derecelik 30 aralığa yerleştirilmektedir ve her pikselin gradyan büyüklük değerinin ilgili açısal bölgedeki değerlere eklenmesi ile oluşturulmuştur. Şekil 4(a)'da görülen yumuşatılmış imgeye uygulanan HOG histogramı Şekil 4(b)'de ve HOG yönelimleri Şekil 4(c)'de verilmiştir. Şekil 5'te bir takım farklı imgelerin histogram dağılımlarına ait grafikler verilmiştir. Bu şekillerden de anlaşılacağı üzere bu HOG değerleri kullanılarak imgenin içerdiği şekil bilgisi elde edilebilir. Kenar sayısının çıkarımında kullanılan kural tabanlı sınıflandırma algoritmasında her bir açısal bölgenin sağındaki ve solundaki komşu açısal bölgeleri ile karşılaştırılmışlardır. Bu karşılaştırmalar bölgelerin farkları alınarak ve bu farkların histogramın ortalamasından büyük olup olmasına bakılarak yapılmaktadır. Bu karşılaştırmalar sonucunda ortalamanın üstünde yer alan her bir bölge bir kenar olarak ele alınmaktadır. Sonuç olarak bulunan kenar sayısı bize geometrik şeklimizin tanımlanmaktadır. Sistemin akış şeması Şekil 3'te verilmektedir.



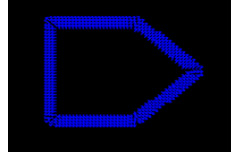
Şekil 3. Sistem akış şeması



(a)

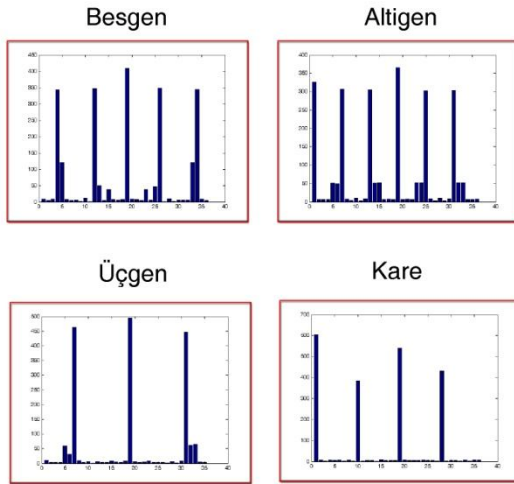


(b)



(c)

Şekil 4. (a) Yumuşatılmış imge (b) HOG projeksiyonu (c) HOG yönelme çizimi



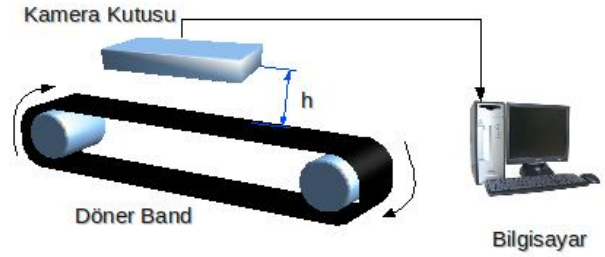
Şekil 5: Farklı şekillere ait gradyan yönleri histogramları

Bu şartlar altında çalıştırılan sistem 125 imge üzerinde %100'lük bir başarımla elde etmektedir.

Sonraki bölümde şekil tanıma sisteminin endüstriyel bir sistem üzerinde gerçek zamanlı uygulaması verilecektir..

4. Endüstriyel Sistem Üzerinde Geometrik Şekillerin Gerçek Zamanlı Tanınması

Bu bölümde endüstriyel sistemde yer alan hareketli bir band üzerinden akmakta olan geometrik şekillerin tanımlanmasına yönelik bir çalışma verilmiştir. Üretim bandı üzerinden geçecek olan geometrik şekillerin web cam ile alınan görüntü üzerinden sınıflandırılması amaçlanmıştır ve donanımdan kaynaklanan sorunların aşılması amaçlanmıştır. Üretim bandı siyah homojen olmayan bir arka plana sahiptir. Görüntü yakalamak için giriş kaynağı olarak kullanılan web cam bir kutu içerisine yerleştirilmiş ve bu şekilde çevresel ışık değişimlerinden etkilenmesi en aza indirilmiştir. Bu band üzerinden geçirilen şekiller ise beyaz renklidir. Sistemin yapısı Şekil 6'te görülmektedir.



Şekil 6. Sistemde kullanılan düzenek

Sistemde bir hareketin var olması dolayısıyla web cam'in görüş alanına bir cismin girip girmediğini algılamak gerekmektedir. Bu algılama işlemi için objenin gri seviyeli histogramında oluşan değişikliklere bakılmıştır. Eğer değişim belirli bir değerin üzerinde ise objenin görüş alanında olduğu anlaşılmaktadır. Bu şekilde yakalanan imgeler (Şekil 7a) ilk olarak 7x7 boyutunda bir gaussian filtresi ile yumuşatılmaktadır (Şekil 7b).



(a)

(b)

Şekil 7: (a) Giriş imgesi (b) yumuşatılmış imge

Yumuşatılan imge (Şekil 7b) kullanılarak yatay ve dikey sobel imgeleri elde edilmektedir (Şekil 8).



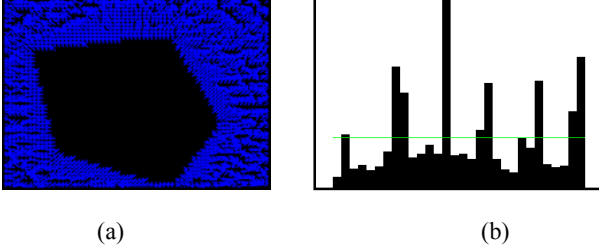
(a)

(b)

Şekil 8: (a) Yatay sobel (b) Dikey sobel

Şekil 8'de ki imgelerden de görülebileceği üzere dikeyde ve yatayda çok fazla gürültü oluşmaktadır. Bu gürültülerin giderilmemesi durumunda çıkarılan HOG histogramı ve HOG

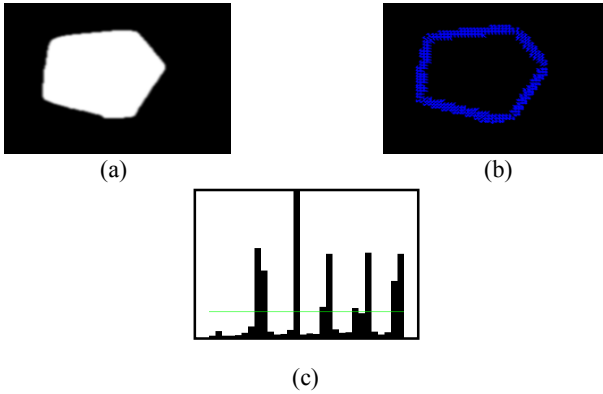
değerlerinin yönelmeleri Şekil 9'da verilmiştir. Şekillerden de görülen bu gürültü ve yönelimlerin sebebi hareketli bandın çizgili bir görünümünün olmasıdır. Bu gürültüler sistemin çalışmasını olumsuz etkilemekte ve sistem bu koşullarda çalıştırıldığı durumda başarımları %46 çıkmaktadır. Sistemin başarımlarının düşük olması aynı zamanda kullanılan kameranın kalitesi ile de alakalıdır.



Şekil 9: (a) HOG yönelimleri (b) HOG değerleri

Bozulmalar ve gürültülerden etkilenecek çalışan sistemin başarımlarının düşük olduğu açıkça görülmektedir. Sistemin başarımlarını düşüren bu gürültü etkenlerinin kaldırılması için yumuşatılmış imge üzerinden bir iyileştirme yapılmaya çalışılmıştır.

Bu iyileştirme yumuşatılmış imgenin (Şekil 7b) piksel piksel taranması ile gerçekleşir. Her pikselin komşu pikselleri ile olan yakınlıkları göz önüne alınır. Komşuluk karşılaştırmaları 3x3'lük bir çerçevede yapılmaktadır. Bu komşuluk oranı hesaplanırken imgenin ortalama değeri sınır olarak kabul edilir. Komşuluk oranının yanında bir başka kriter olarak pikselin beyaz renge yakınlığı alınmaktadır. Ortalama değerin üzerindeki pikseller beyaz, bu değeri geçemeyen pikseller ise siyah olarak işaretlenmektedir. Bu işaretlemeler tamamlandıktan sonra objenin bulunduğu bölge 5 piksel genişletilerek Şekil 10a'da görülen imge elde edilmektedir.



Şekil 10: (a) Maske imgesi (b) Maskelenmiş HOG yönelimleri (c) Maskelenmiş HOG değerleri

Şekil 10a'da görülen maske, HOG çıkarımı yapılan algoritmaya ek parametre olarak verilmektedir. Maskelenmiş piksellerde HOG hesaplaması yapılmamaktadır. Böylece işlem yükü azaltılmaktadır. Maskelenmiş HOG sonucunda elde edilen imge Şekil 9b'de görülmektedir. Şekil 9a ile Şekil 10b ve Şekil 9b ile Şekil 10c karşılaştırıldığı zaman görülecektir ki gürültü öğeleri oldukça azaltılmıştır. Böylece hem gürültü etkenleri hem de hız açısından iyileşme elde edilmektedir.

Ayrıca bir nesnenin kamera önünden geçiş süreci tamamlandığı zaman o nesnenin geçişi esnasında yapılan bütün tanıma sonuçları birleştirilmektedir. Bu sonuçlar değerlendirilirken oylama sistemi ile en yüksek puanlı kenar sayısı sonuç olarak kabul edilmektedir. Bu iyileştirmeler sonucunda sistemin başarımları %91,41'e çekilmiştir. Ayrıca yapılan iyileştirme ile %23'lük bir hız artışı elde edilmektedir.

5. SONUÇ

HOG ile şekil tanıma ve endüstriyel bir uygulamada karşılaşılan problemlere getirilen çözümler bu çalışma ile verilmiştir. Gürültü faktörlerinin temizlenmesi ile sistem başarımlarının artırıldığı görülmüş ve HOG tanımlayıcısı ile yapılan algoritmanın gerçek zamanlı bir sistem için iyileştirilmesi yapılmıştır.

Bu çalışma da bir sınıflandırıcı kullanılması ve kullanılan kamera ve düzeneğin kaliteli olması durumunda başarımların ideal ortamdaki başarımlara ulaşması çok muhtemel bir durumdur. Bu uygulama ile HOG kullanılarak yapılabilecek tanımlama işlemlerine bir örnektir.

6. Teşekkür

Bu uygulama SAN-TEZ 168STZ2007-2 projesi kapsamında Sanayi ve Ticaret Bakanlığı tarafından desteklenmiştir.

6. Kaynakça

- [1] Shashua A., Gdalyahu Y., and Hayon G., "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance", In Proceedings of IEEE Intelligent Vehicles Symposium, 2004.
- [2] N. Dalal and B. Triggs., "Histograms of oriented gradients for human detection", In C. Schmid, S. Soatto, and C. Tomasi, editors, International Conference on Computer Vision and Pattern Recognition, volume 2, pages 886-893, June 2005.
- [3] Bay, H., Tuytelaars, T., and Gool, L. V. (2006). Surf: "Speeded up robust features", In Leonardis, 2006.
- [4] Zhu Q., Mei-Chen Y., Kwang-Ting C., Avidan, S., "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients", Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, 2006.
- [5] Ong E., Bowden R., "A Boosted Classifier Tree for Hand Shape Detection", Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004.
- [6] D.M. Gavrila, V. Philomin, "Real-Time Object Detection For 'Smart' Vehicles", Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference, 1999.
- [7] Bahlmann C., Zhu Y., Ramesh V., Pellkofert M., Koehled T., "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information", IEEE, 2005.
- [8] Miura J., Kanda T., Shirai Y., "An Active Vision System for Real-Time Traffic Sign Recognition", 2000 IEEE Intelligent Transportation Systems Conference Proceedings, October 2000.

- [9] Loy G., Barnes N., "Fast shape-based road sign detection for a driver assistance system, Intelligent Robots and Systems", 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference, 28 Sept.-2 Oct. 2004.
- [10] A. Broggi, M. Bertozzi, A. Fascioli, M. Sechi, "Shape-based Pedestrian Detection", IEEE Intelligent Vehicles Symp, 2000.
- [11] Rong-Chin L., Wen-Hsiang T., "Perspective-Transformation-Invariant Generalized Hough Transform For Perspective Planar Shape Detection and Matching", Pattern Recognition, Vol 30, No. 3, pp. 383-396, 1997.