Market Raflarında Dönüşüm ve Ölçeklendirmeye Dayanıklı Nesne Tanıma

Melih Evren Buruş[1], Rıdvan Salih Kuzu[2,3], Soyhan Beyazıt[1,3], Gül Varol[1,3]

[1] Boğaziçi Üniversitesi, Bilgisayar Mühendisliği Bölümü, İstanbul

[2] Boğaziçi Üniversitesi, Elektrik ve Elektronik Mühendisliği Bölümü, İstanbul

[3] İdea Teknoloji Çözümleri Bilgisayar San. Ve Tic. Ltd. Şti, İstanbul

melih.burus@boun.edu.tr, ridvan.kuzu@ideateknoloji.com.tr, soyhan.beyazit@ideateknoloji.com.tr, gul.varol@ideateknoloji.com.tr

Özet: Son kullanıcılar hangi ürünü satın alacaklarına çoğunlukla marketlerde karar vermektedirler. Üreticilerin başarısı üretim ve pazarlama güçlerini tedarik zincirinin son noktasına yansıtmaları ile ilişkilidir. Üreticiler ürünlerinin marketlerde konumlandırılma şeklini takip edebilmek için kullanımı kolay bilgi toplama araçlarına ihtiyaç duymaktadırlar. Bu çalışmada varolan nesne tanıma teknikleri araştırılmış, market raflarında ürünlerin planogram bilgisi göz önünde bulundurularak marka bazında doğru konumlandırılıp konumlandırılmadıklarını tespit edebilen bir sistem önerilmiştir.

Anahtar Sözcükler: Öznitelik Tespiti, Köşe Tespiti, Nesne Tespiti, Nesne Tanıma, Afin Değişmezlik.

Rotation and Scale Proof Object Recognition for Market Shelves

Abstract: Today's most of the purchasing decisions are made in retail market. Reflecting manufacturers' production and marketing power to the end-point is an important aspect for their success. In order to track positioning of their goods in the retail market, manufacturers need to have easy-to-use information collecting tools. In this preliminary work, we will study existing object recognition techniques in the literature and propose a system that will be able to detect the goods in brand specific manner on market shelves and to decide whether they are on the right positioning or not, according to their planogram information.

Keywords: Feature Detection, Corner Detection, Object Detection, Object Recognition, Affine Invariance.

1. Giriş

Nesne tanıma bilgisayarla görme alanının bir alt dalı olup sayısal imge ve videolarda nesnelerin ayırt edici özelliklerine göre teşhis edilmesi olarak tanımlanmaktadır. İnsanlarda görme sisteminin tanımlanması ve bu alandaki gelişmeler sonrasında bilgisayarla görme ve nesne tanıma sistemleri de evrilmiştir. Bilgisayarla görme konusunda çeşitli yöntemler geliştirilmiş, bu yöntemler yüz tanıma, iris tanıma, optik karakter tanıma, vücut ve aktivite takibi, web tabanlı sistemlerde imge araması, barkod tanıma, tıbbi görüntüleme ve oyunlarda görme-tabanlı etkileşim gibi uygulama alanlarında uygulanmış ve araştırmalara halen devam edilmektedir.

Bir bilgisayarla görme sistemi temel olarak önişleme, segmentasyon, öznitelik çıkarımı ve sınıflandırma adımlarından oluşmaktadır. Sırası ile imge filtrelenip geliştirilir, nesneler arka alandan ayrıştırılır, her nesnenin ayırt edici özellikleri çıkarılır ve nesneler önceden tanımlanmış olan sınıf ya da kategorilere dahil edilir.

Nesnelerde ideal bir öznitelik perdelemelere, karışmalara, parazitlere ve bulanıklıklara dayanıklı iken aynı zamanda değişmez, ayırt edici, doğru ve verimli çalışır olmalıdır. Nesne tanıma tekniklerinin dayanıklılığı nesnelerin dönüşüm ve ölçeklendirmeye karşı değişmez özniteliklerini çıkarabilmesine bağlıdır.

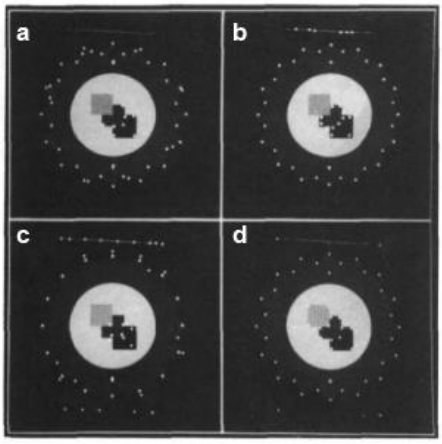
Sonuç olarak görüntünün analizi, ideal özniteliklerin çıkarımı ve planogram ile eşleştirilmesi önerilen sistemin en önemli bileşenleri olarak öne çıkmaktadır. Bu amaçla bilinen tekniklerin (köşe sezinleyiciler, kenar sezinleyiciler, Gaus türevi, Laplas türevi, LoG operatörü, HOG tanımlayıcısı vb.) nasıl çalıştığının anlaşılması uygun yöntemi belirlemek için önemlidir.

Bölüm II'de Moravec ve Harris köşe sezinleyicileri anlatılmaktadır. SIFT (Ölçek Bağımsız Öznitelik Dönüşümü), Harris-Afin ve Hessian-Afin ve SURF yöntemleri afin değişmez öznitelik sezinleyicileri olarak Bölüm III-VI'da anlatılmaktadır. Bölüm VII'de HOG tanımlayıcıları, Bölüm VIII'de ise ölçek ve afin değişmez algılayıcıların performans ve doğruluk kriterlerine göre farklı koşullarda tercih edilme sebepleri ve çözüm önerimiz anlatılmıştır.

2. Köşe Sezinleyiciler

2.1 Moravec Köşe Sezinleyicisi

Moravec köşe sezinleyicisi, ilk köşe sezinleme algoritmalarından biridir ve her pikselin yakınında örtüştüğü pencereleri bulup kare farkı algoritmasını kullanarak pencereler arasındaki parlaklık farkını hesaplar [1]. Parlaklık değişimi sabite yakınsa değişiklikler çok küçüktür. Piksel kenara denk geliyorsa kenar doğrultusunda ilerledikçe çok küçük değişiklikler, kenara dik yönde ilerledikçe büyük değişiklikler beklenir. Piksel bir köşe ya da izole olmuş nokta ise bütün yönlerde yapılan ilerlemelerin büyük değişikliklere sebep olması beklenmektedir. Şekil 1(a)'da açık renkli noktalar Moravec Köşe Sezinleyicisi tarafından bulunan köşeleri temsil etmektedir.



Şekil 1 (a) Moravec (b) Beaudet (c) Kitchen ve Rosenfeld (d) Harris [2].

2.2 Harris Köşe Sezinleyicisi

Harris ve Stephens 1988'de elde ettikleri sonuçlara göre iyi bir imge özniteliğinin ayrık olması gerektiğini bu yüzden kenar sezinleyiciler (Canny) yerine köşe sezinleyicilerin kullanılabileceğini belirtmiş, değişken pencereleri kullanmak yerine köşe skorunun (Otokorelasyon-Harris Matrisi) türevini alarak Moravec köşe sezinleyicisini de geliştirmişlerdir.

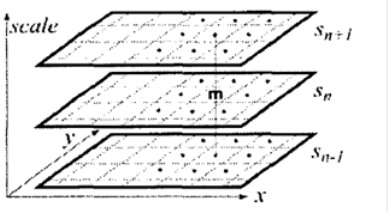
MCD dikdörtgen şeklinde ikili değerlerden oluşan imge pencerelerini kullanmaktadır. Harris sezinleyicisi ise Gaus fonksiyonu (yumuşatılmış ve dairesel) kullanmaktadır. Parlaklık farkı kümesinin sadece minimum değerleri göz önüne alındığı için MCD'nin kenarlara tepkisi fazladır. Lokal maksimumu aramak yerine hareket yönünde gerçekleşen varyasyonu gözlemlemek kenarlara olan tepkisini düşürecektir. Şekil 1(d) Harris Sezinleyicisi tarafından bulunan köşeleri göstermektedir [2].

Parlaklık değişimi Harris matrisi olarak da adlandırılan otokorelasyon matrisi ile temsil edilir. Köşe sezimi (lokal maksimum) özdeğerlerin dekompozisyonunu gerçekleştirmek yerine matrisin izi ve determinantı kullanılarak hesaplanmıştır. Harris özdeğerlerini göz önünde bulundurduğumuzda; ikisi de birden küçükse imge penceresi parlaklık olarak neredeyse sabittir. Yüksek-düşük ya da düşük-yüksek eğrilik çiftleri bir kenarın bulunduğu ve kenar boyunca küçük, kenara dik yönde ise büyük değişimlerin olduğu anlamına gelir. Her ikisinin de yüksek olduğu durumlarda köşe olduğu anlamına gelmektedir

3 Ölçek Bağımsız Öznitelik Dönüşümü (SIFT)

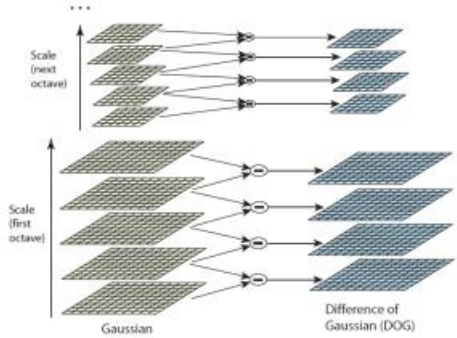
İdeal bir nesne özniteliği ölçekten bağımsız olmalıdır, bu yüzden farklı ölçeklerde belirlenmesi ve tanımlanması mümkün olmalıdır. Ölçekten bağımsız öznitelik çıkarımında kullanılan ilk yöntemler imgelerin 3B ölçek uzayında (x,y ve ölçek) temsili ve lokal maksimum arayışına dayanmaktadır. Bu maksatla Lindeberg Gaus'un Laplas fonksiyonunu (LoG), Lowe ise Gaus'un türevini (DoG) kullanmayı önermiştir. Şekil 2 bir imgenin x y ve ölçek faktöründen oluşan ölçek uzayda temsilini göstermektedir [6].

Lindeberg LoG fonksiyonunu kullanarak imgeleri 3B ölçek uzayında temsil etmiştir. Çalışmasındaki temel fikir imge özniteliklerini belirleyip Gaus türev filtrelerine dayalı bir ölçek seçme prensipi oluşturmaktır. Buna göre farklı ölçeklerde lokal maksimum aranarak bir özniteliğin ne kadar büyük olduğuna karar verilir. Öznitelik tespitinin kaba ölçeklerde, ana nokta konumlandırmasının ise ince ölçeklerde yapılması gerekir. LoG farklı öznitelik algılama yöntemleri ile birlikte kullanılabilmektedir (blob sezimi, köşe sezimi, kenar sezimi gibi) [3].



Şekil 2 Ölçek uzayında maksimum arama [6].

1999 yılında Lowe, DoG ile ölçek uzayın minimum ve maksimum değerlerini çıkararak LoG fonksiyonunu hesaplamış ve yöntemi Ölçek Bağımsız Öznitelik Dönüşümü (SIFT) olarak adlandırmıştır. SIFT, ölçek uzayda DoG (Şekil 3) fonksiyonunun minimum ve maksimum değerlerini belirleyerek dönüşümden bağımsız ve daha verimli öznitelik tespiti gerçekleştirir [4].



Şekil 3 DoG fonksiyonu ile imgelerin ölçek uzay temsili [7].

SIFT, piramidin her seviyesinde piksellerin gradyan boyunu ve yönünü hesaplar. Işıklandırma farklılıklarına karşı dayanıklılık için gradyan boyları maksimum gradyan değerinin 0.1 katı kullanılarak eşiklendirilir. Ana noktalar kanonik oyantasyonda atandığı için imge tanımlayıcıları dönüşümden bağımsız olarak elde edilir [4].

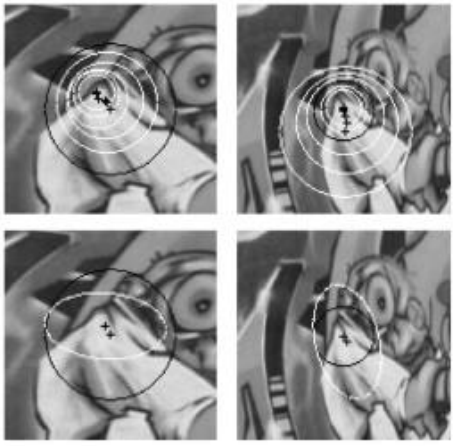
Ölçek bağımsız bir diğer yöntem olan Harris-Laplace Sezinleyicisi, ölçek bağımsız ilgi noktaları elde etmek için Harris köşe sezinleyicisini uygulayarak Gaus ölçek uzay temsilini oluşturur. Harris-Laplace öncelikle Gaus fonksiyonu ile yumuşatılmış olan imgeden Harris Matrisi olarak da bilinen ikinci derece moment matrisini hesaplar.

Köşeyi temsil eden lokal maksimum değeri Harris Matrisi'nin determinant ve izi kullanılarak hesaplanabilir. Bu noktada Harris-Laplace Sezinleyicisi ölçek uzay noktalarını belirler ve LoG fonksiyonunun maksimum değer aldığı noktaları seçer [6].

4 Harris-Afin ve Hessian-Afin Sezinleyicisi

Mikolajczyk ve Schmid, 2002 yılında Harris-Laplace tabanlı afin değişmez bir yöntem önerdiler. Dönüşüm, ölçek ve çevirmeye dayanıklı ilgi noktalarını tespit edebilmek için Harris-Laplace sezinleyicisinin çok ölçekli gösterimini gerçekleştirip türevinin lokal maksimumunu hesaplattılar.

Afin dönüşümlere farklı ölçeklerde bakılırsa her yönde aynı oranda gerçekleşmediği görülür (Şekil 5). Bu yüzden seçilen ölçek gerçek afin dönüşüm noktasını yansıtmayabilir. Bu problemin üstesinden gelmek için afin Gaus ölçek uzayında ikinci moment matrisi kullanılmıştır. Dairesel pencereler kullanmak yerine eliptik pencere kullanımına denk gelen afin Gaus ölçek uzayı düzensiz Gaus çekirdekleri kullanılarak hesaplanabilmektedir [8].



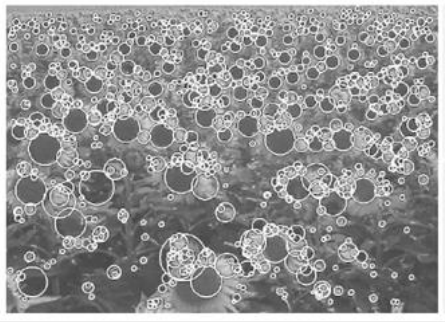
Şekil 2. Üst sıra Harris-Laplace, alt sıra afin dayanıklı Harris-Laplace sezinleyicisi [10].

Beaudet 1978 yılında yaptığı çalışmada imgelerin Hessian matris determinantlarını hesaplayarak köşe noktalarının lokal uç noktaları olarak bulunabileceğini göstermiştir [9].

Hessian-Afin ile Harris-Afin sezinleyicileri ikinci moment matrisi olarak Harris yerine Hessian matrisi kullanmak dışında hemen hemen aynıdır. DoG ve LoG uzaylarında düz kenarlara ve sınırlara yakın pikseller lokal maksimum olarak algılanabilmektedir. Bu problem Hessian matrisinin determinant ve izi kullanılarak çözülmüştür.

5 SURF Sezinleyicisi

Bay, ölçek ve ilgi noktaları seçiminde Hessian matrisinin determinant ve izini (Laplas) kullanmak yerine, sadece determinantını hesaplamayı önermiştir. SURF metodu daha iyi bir performans sağlamak için iki yöntemi birleştirerek yeni bir yaklaşım ortaya koymuştur. Hessian matrisi ve basit Laplace hesaplamasına dayananan algoritma Fast-Hessian sezinleyicisi olarak da bilinir. SIFT ile benzerlikler gösteren tanımlama algoritması, ilgi noktalarının Haar-Wavelet algoritmasına tepkisini bulur ve daha düşük işlemsel karmaşıklık değerine sahiptir. Şekil 6, ayçiçeği tarlasından alınan bir imgeden Fast-Hessian algılayıcısı ile çıkartılan ilgi noktalarını gösterir [11].



Şekil 3 SURF ile ayçiçeği tarlasından tespit edilen ilgi noktaları [11].

6. HOG Tanımlayıcıları

Çakışan imge bloklarından elde edilen gradyan değerlerinin histogramının çıkarılması ile öznitelik tanımlama ilk kez

Dalal ve Triggs tarafından ortaya konulmuştur [12]. İlk kez yaya tanıma alanında kullanılan ve HOG olarak adlandırılan bu yöntem daha sonra farklı nesne tanıma problemlerinde de yaygınlık kazanmıştır.

Bu yöntemde, imgedeki bir nesneyi bulmak için, aranan nesneyle aynı boyutta pencere kullanılır. Pencere belirli bir çakışma oranıyla imgenin üzerinde kayar, her bir pencere bloğunun kendi içinde çakışan hücrelere bölünmesi ile hücrelerin üzerine geldiği her alanda lokal öznitelikleri tespit eder.

HOG özniteliklerinin elde edilmesinde kullanılan temel mantık, lokal imge hücrelerinin her birinden elde edilen gradyan ve sınır yönelim doğrultularının yoğunluk dağılımına ulaşılmasıdır. Yoğunluk dağılımı için kullanılan en yaygın yöntem, her bir yönelim doğrultu açısının tespit edilmesi ve 360 derecenin belirli kutulara bölünmesiyle, gelen her açı değeri için o kutudaki değerin artırılmasına dayanır.

7. Uygulama Çözüm Önerimiz

Proje sonunda elde edilmek istenen sistemin, raflardaki ürünlerin saha dizilimiyle, planogramda tanımlı beklenen dizlimin uyumlu olup olmadığını tespit edebilmesi hedeflenmektedir. Bu amaçla yürütülen çalışmalarda yatay doğrultuda türevlenmiş raf imgelerinin Hough Doğrultu Tanıma algoritması ile yatay çizgileri Şekil 7'de de görüldüğü gibi çıkarılmıştır.

İkinci adımda her bir raf aralığı ayrı ayrı ele alınarak, ilgili raftaki nesnelerin tespit edilmesi aşamasına geçilmiştir. Pozitif veri kümesi (üzerinde ürün olan imgeler) ve negatif veri kümesi (üzerinde ürün olmayan imgeler) için ayrı ayrı elde edilen HOG öznitelikleri Kademeli Nesne Tanıma sınıflandırıcısı ile eğitilmiş [13] ve nesnelerin raflar üzerinde konumları Şekil 8'de görüldüğü gibi belirlenmiştir.



Şekil 7: Raf yatay çizgilerinin tespiti

Uygulamanın son aşaması olarak da yerleri tespit edilen her bir nesnenin kendi sınırları içinde marka ve ürüne ait özniteliklerin çıkarılması planlanmaktadır. Planogram bilgisine dayanılarak rafın ilgili yerinde olması gereken ürünün öznitelikleri ile gerçekte olan ürünün öznitelikleri karşılaştırılarak bire-bir doğrulama çalışması yapılacak ve rafların planograma uygun dizilmemiş yerleri varsa tespit edilebilecektir.



Şekil 8: Raftaki nesnelerin yerlerinin tespiti

8. Sonuç ve Öneriler

Yöntemlerin performans ve doğruluk değerleri uygulamaların sezinleme yanında tanımlama ve eşleştirme bileşenlerine de bağlıdır. Bu yüzden veri kümesinin türüne ve ortam kısıtlarına göre en uygun yöntem değişiklik göstermektedir. Ayrıca, öznitelik sezinleyicilerinin işlemsel karmaşıklığı ve doğruluğu arasında ödünleşim vardır.

Hessian tabanlı sezinleyiciler Harris tabanlı sezinleyicilere göre daha tutarlı ve tekrarlanabilirdir. DoG tabanlı sezinleyiciler işlemsel performans açısından daha iyi çalışmakla birlikte düşük de olsa doğruluktan ödün verilmektedir. SIFT, yaygın olarak kullanılan, hızlı ve ayırt edici bir yöntemdir. HOG, imgelerin lokal kısımlarında sabit boyutlu nesnelerde yüksek işlemsel performans sunan bir yöntemdir.

Uygulamada ortaya konulan çözüm önerisinde, rafların sınır tespitinde kullanılan yöntem 229 raf imgesi üzerinde test edilmiş, bir resimdeki tüm rafların bulunmasının doğru, tek bir rafın bile kaçırılmış olmasının yanlış kabul edildiği yaklaşımda % 83.4 oranı ile raflar doğru tespit edilmiştir. Ayrıca 291 ürün konumu için yapılan nesne tespiti testlerinde ise *precision* ve *recall* testleri sırasıyla % 88.8 % 87.9 sonuçlarını vermiştir.

9. Teşekkür

Bu çalışma İdea Teknoloji ve Bilgisayar Çözümleri Ltd. Şti firmasının TEYDEB 1501 projesi kapsamında yapılmıştır. Ayrıca, Boğaziçi Üniversitesi Bilgisayar Mühendisliği Bölümü'nden Prof.Dr. Fatih Alagöz ve G.Y.T.E. Bilgisayar Mühendisliği Bölümü'nden Doç.Dr. Yusuf Sinan Akgül'e katkılarından dolayı teşekkür ederiz.

10. Kaynaklar

[1] H. Moravec, Obstacle Avoidance and navigation in the real world by a seeing robot rover, tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University & doctoral dissertation, Stanford University, September, 1980.

[2] C. Harris, M. Stephens, A combined corner and edge detector, in Proceedings of the Fourth Alvey Vision Conference, Vol. 15, pp. 147–151, 1988.

[3] T. Lindeberg, Feature detection with automatic scale selection, International Journal of Computer Vision, vol 30, number 2, pp 77--116, 1998.

[4] D. Lowe, Object recognition from local scale-invariant features, Proceedings of the International Conference on Computer Vision. 2. pp. 1150–1157, 1999.

[5] D. Lowe, Distinctive image features from scale-invariant key points, Int. J. Comput. Vis., 60, pp. 91–110, 2004.

[6] K. Mikolajczyk, C. Schmid. Indexing based on scale invariant interest points. In Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada, pages 525–531, 2001.

[7] K. Mikolajczyk, C. Schmid, scale & affine invariant interest point detectors. International Journal on Computer Vision 60(1):63–86, 2004.

[8] K. Mikolajczyk, C. Schmid, An affine invariant interest point detector, Proc. Seventh European Conf. Computer Vision, pp. 128-142, 2002.

[9] P.R. Beaudet, Rotationally invariant image operators, Proceedings of the International Joint Conference on Pattern Recognition, Kyoto, Japan, pp. 579–583, 1978.

[10] K. Mikolajczyk, C. Schmid, Scale and affine invariant interest point detectors, IJCV 60, 63–86, 2004.

[11] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded up robust features. In ECCV (1), pages 404-417, 2006.

[12] N. Dalal, B. Triggs, I. Rhone-Alps, and F. Montbonnot, "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005

[13] J. Viola, J. Paul, J. Michael. "Rapid object detection using a boosted cascade of simple features." IEEEComputer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2001