

Rafta Ürün Diziliminin Görüntü İşleme ile Tespiti

Product Placement Detection Based on Image Processing

Gül Varol^{*†}, Rıdvan Salih Kuzu^{†‡}, Yusuf Sinan Akgül[§]

^{*}Bilgisayar Mühendisliği Bölümü, Boğaziçi Üniversitesi

[†]Elektrik ve Elektronik Mühendisliği Bölümü, Boğaziçi Üniversitesi

[‡]İdea Teknoloji Çözümleri

[§]Bilgisayar Mühendisliği Bölümü, Gebze Yüksek Teknoloji Enstitüsü
{gul.varol, ridvan.salih}@boun.edu.tr, akgul@bilmuh.gyte.edu.tr

Özetçe —Marketlerde ürünlerin konumlandırılma şekli üreticiler için önem taşımaktadır. Bu yüzden raflardaki dizilim planogram ismi verilen belgelerle yapılandırılır. Ürün yerleştirmesinin planograma göre yapıp yapılmadığını takip edebilmek için üreticiler kolay bilgi toplama araçlarına ihtiyaç duymaktadırlar. Bu çalışmada, var olan görüntü işleme teknikleri kullanılarak, marketlerdeki planogram uyumluluğu kontrolünün otomatikleştirilmesi amaçlanmıştır. Planogram bilgisi bağlam olarak kullanılarak market görüntülerinden anlamlı bilgi çıkartılabileceği varsayımıyla envanter yönetimi için yenilikçi bir yöntem önerilmiştir. Görevli kişi, rafların fotoğrafını çekecek ve sistemimiz tarafından işlenmek üzere görüntüyü sunucuya yükleyecektir. Sistem, üç ayrı birimle çalışmaktadır: 1) rafların tespiti, 2) ürünlerin tespiti ve 3) marka tanıma. Temel olarak raf tespitinde Hough Dönüşümü'ne, ürün tespitinde Kademeli Nesne Tanıma algoritmasına, ve marka tanımda Destek Vektör Makinesi'ne başvurulmuştur. Sistemin sınanması için sigara ürünlerinden görüntü veri kümesi oluşturulmuş, her birim ayrıca sınanmıştır.

Anahtar Kelimeler—Nesne Tespiti, Nesne Tanıma

Abstract—Product positioning in stores is of importance for the manufacturers. With this aim, the placement on shelves is done according to some documents named planograms. To check whether the placement is in compliance with the planogram, manufacturers need tools to gather information easily. In this work, using existing image processing techniques, we aim for the automation of planogram compliance control. We propose a novel technique for inventory management with the assumption that we can extract meaningful information from images using planogram context. The controller will be able to capture a photograph of the shelves and upload to the server to be processed by our system. The system is composed of three modules: 1) shelf detection, 2) product detection and 3) brand recognition. Basically, we make use of Hough Transform, Cascade Object Detection algorithm, and Support Vector Machines for the modules respectively. We test each module separately on a dataset we collected with cigarette products.

Keywords—Object Detection, Object Recognition

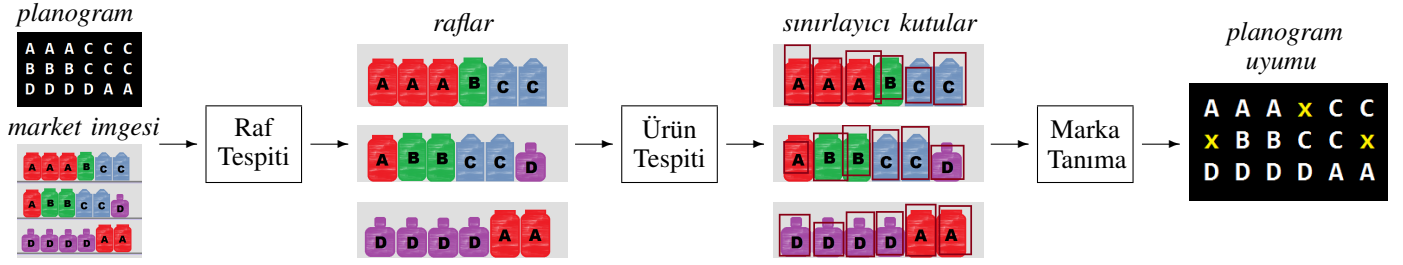
I. GİRİŞ

Planogramlar, perakende ürünlerin marketlerdeki yerleştirmelerini belirten şemalardır ve birtakım pazarlama stratejilerine ve kurallarına göre tasarlanmışlardır. Günümüzde market raflarının bu şemadaki gösterime uyup uymadığı belirli aralıklarla gönderilen görevliler tarafından kontrol edilir. Çoğunlukla, görevli kişi, ürünleri manuel olarak sayar veya daha sonra yine manuel olarak kontrol edilmek üzere ürünlerin fotoğrafını çeker. Son zamanlarda bu kontrolün otomatikleştirilmesi için çalışmalar yapılmıştır [1], [2]. Planogram uyumluluğunun otomasyonu daha hızlı ve güvenilir olduğundan envanter yönetimi için büyük önem taşır.

Literatürde görüntü işlemeyle planogram çıkartılması için SIFT (scale-invariant feature transform) [3] ve optik karakter tanıma gibi yöntemlerin birleşimini kullanan bir patent [4] bulunmaktadır. Bunların yanı sıra, planogram çıkartılması problemine benzer problemler için çeşitli yöntemler önerilmiştir. [5]'te doğal sahnelerde logo tespiti için uzamsal piramit madenciliği ve SIFT tanımlayıcıları kullanılmıştır. [6]'da kütüphanedeki kitapları tanıma üzerine bir yöntem sunulmuştur, ancak buradaki yaklaşım bizim problemimizdeki görüntüler için yetersizdir.

Market ortamındaki görüntülerle çalışan bir diğer çalışma [7]'de sunulmuştur. Bu makalede, görme engellilerin alışveriş listelerindeki ürünleri kamera yardımıyla bulması hedeflenmiştir. Gerçek zamanlı çalışma kısıtı bulunduğundan, daha hızlı hesaplanan SURF (speeded up robust features) tanımlayıcıları tercih edilmiştir. Bir başka çalışmada ise bir girdi görüntüde, veritabanındaki imgelerden hangilerinin bulunduğunu SIFT tanımlayıcıları yardımıyla bulma problemi ele alınmıştır [8]. Bu çalışmada market ve ürün imgeleriyle uygulamalar yapılmıştır; ayrıca yavaş çalışan sistemin hızlandırılması için eniyileme çalışmaları yapılmıştır.

Bildirinin devamı şu şekilde düzenlenmiştir. II. bölümde yöntem sunulmuştur. Sistemin hangi alt birimlere nasıl ve neden ayrıştırıldığı açıklanmıştır. III. bölümde her bir birimi sınamak amacıyla uygulanan deneyler ve sonuçları sunulmuştur. Son olarak IV. bölümde alınan sonuçlar değerlendirilmiştir ve bu çalışmanın nasıl geliştirilebileceği ile ilgili birkaç nokta belirtilmiştir.



Şekil 1: Sistemin genel akışı.

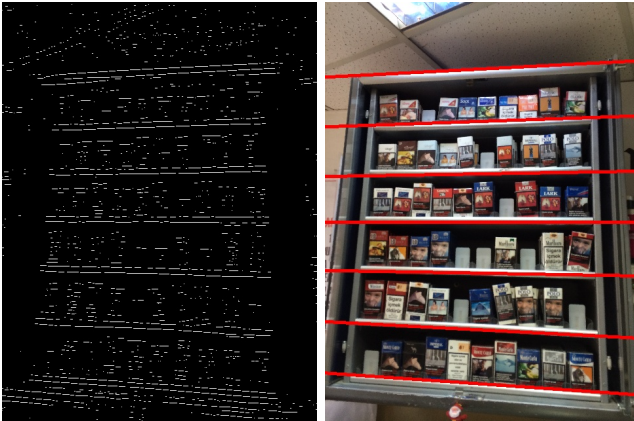
II. YÖNTEM

Bu çalışmada önerdiğimiz yöntemde market ürünlerinin güvenilir bir şekilde tanınması için üç aşamalı bir işlem uyguluyoruz. Bir ürünün logosunu bütün imgede aramak yerine, herhangi bir nesne tanıma gerçekleştirmeden önce rafların yapısı ve ürünlerin dağılımı bilgisini imgeden çıkartıyoruz.

Önerdiğimiz yöntemde, ürünlerin matrise benzer bir yapıda dizilime sahip olduğunu varsayıyoruz. Bu yüzden ilk aşamada bu ızgaranın sınırlarını belirlemek istiyoruz. Bu amaçla raf tespiti, ardından ürün ve marka tespiti uyguluyoruz. Şekil 1’de gösterilen akışta sistemin genel mimarisi, girdi ve çıktıları özetlenmiştir.

A. Raf Tespiti

Raf sınırlarını tespit ederek arama alanını kısıtlamak ve imgeyi anlamlı bölgelere ayırmak hedeflenmiştir. Bu amaçla yatay doğruları bulmak için Hough Dönüşümü uygulanmaktadır [9]. Bu dönüşümün kullanılabilir sonuç verebilmesi için imge birtakım ön işlemlerden geçirilir. Bunlar, imgeyi gürültüden arındırmak için Gaussian bulanıklığı uygulama; kenarları vurgulamak için Canny kenar belirleme algoritması uygulama; yataydaki çizgileri öne çıkarmak için yatayda kenar bulma çekirdeği ile filtreleme; ve bir eşik değeri ile imgeyi ikili görüntüye çevirme olarak özetlenebilir. Ayrıca sistemi parametrelerden olabildiğince bağımsızlaştırmak için, ilk olarak her imge aynı yüksekliğe (640 piksel) sahip olacak şekilde en-boy oranı bozulmadan yeniden boyutlandırılır. Şekil 2 örnek bir imge üzerinde bahsedilen işlemlerden oluşan çıktıyı görselleştirmektedir.



Şekil 2: Ön işlemin çıktısı (solda). Raf sınırları (sağda).

Ön işleme uygulanmış imge üzerinde Hough Dönüşümü uygulanır, fakat raf sayısından çok daha fazla yatay doğru tespit edildiği gözlemlenmiştir. Bu doğrular arasında raf sınırlarına denk gelenleri seçmek için yukarıdan aşağı (top-down) bir yaklaşım uyguluyoruz. Eşleştirilecek planogramda raf sayısı bilgisi bulunmaktadır; dolayısıyla bulunması gereken raf sayısı sisteme girdi olarak verilebilir. Toplamda tespit edilen doğru sayısı N , raf sayısı ise n olursa ($n \leq N$); N ’in $(n + 1)$ ’li kombinasyonları içerisinde bir arama yapıp en uygun doğru kombinasyonu raf sınırları olarak seçilebilir. Bu seçimi yaparken kullandığımız kriter, rafların birbirine eşit uzaklıkta/eşit uzaklığa yakın bulunduğu yönündedir. Her $C(N, n + 1)$ kombinasyonu için Denklem (1)’de ortalama kare hatası olarak tanımlanan enerji fonksiyonu hesaplanır.

$$E_c = \sum_{\ell \in C} (\rho_\ell - \hat{\rho}_c)^2 \quad (1)$$

Burada C , o anki doğruların kümesi ve ρ_ℓ ise ℓ doğrusu ile orijin (imgenin sol üst köşesi) arasındaki uzaklığı temsil eder. Ortalama $\hat{\rho}_c$ ise şu şekilde hesaplanır:

$$\hat{\rho}_c = \frac{\sum_{\ell \in C} \rho_\ell}{|C|} \quad (2)$$

Bu tanımlardan yola çıkarak bir C kombinasyonu için enerji değeri E_c ’nin sıfıra yakın olması eşit mesafe özelliğini barındırdığının göstergesi olur. Dolayısıyla aranılan doğru kümesi C^* , minimum enerji değerine sahip olan olarak tanımlanır:

$$C^* = \arg \min_C E_c \quad (3)$$

Bazı durumlarda C^* çözümündeki doğrular birbirine çok yakın doğrulardan çıkabilir. Bunun sebeplerinden bir tanesi Hough Dönüşümü’nün sonucunda tek bir kenarın etrafında yoğunlaşmış, çok sayıda doğrunun tespit edilmesidir. Bu sorunu basit bir eşikleme uygulayarak çözüyoruz.

$$t = \frac{h}{2n} \quad (4)$$

Burada h imge yüksekliğini, t ise eşik değerini göstermektedir. Her ℓ için $\rho_\ell - \hat{\rho}_c > t$ kısıtını sağlayan kombinasyonlar çözüm için ele alınır. Böylece birbirine çok yakın, fakat eşit uzaklık özelliği gösteren kombinasyonlar elenmiş olur. Dolayısıyla n tane raf, imgenin dikey ekseninde en az yarısını kaplar varsayımı



Şekil 3: Tüm doğrular $N = 50$ (solda), en uygun kombinasyon $n = 3$ (sağda).

kullanılır. Bu geçerli bir varsayımdır, çünkü fotoğraf, raflar çerçeveye tam sığacak şekilde yakından çekilmelidir.

Ürünlerin konumları için kısıtlamalar getiren raf tespit birimi, çıktı olarak raf bölgelerini vermektedir. Raf sınır çizgilerinden imge bölünür ve her bir raf aralığı için ürün tespiti birimi uygulanır. Örnek bir raf tespiti imgesi Şekil 3'te gösterilmiştir.

B. Ürün Tespiti

Ürün tespiti birimi için Viola-Jones nesne tanıma iskeletini kullandık. İlk olarak yüz tespit etme problemi için geliştirilen bu iskelet, AdaBoost öğrenme algoritması ile kademeli bir sınıflandırma yapısı içerir [10]. En-boy oranı çok değişmeyen, düzlem dışı dönme içermeyen nesnelerin tespiti için bu yöntemin iyi çalıştığı bilinmektedir.

Bu çalışmada öznitelik olarak HOG (Histogram of Oriented Gradients) tanımlayıcılarını kullanıyoruz, ve kendi ürün sınıflımız için bir sınıflandırıcı eğitiyoruz. Marka farkı gözetmeksizin ürün imgeleri pozitif, ürünün tamamını içermeyen (arkaplan, raf sınırları, iki ürünün arası vb.) imgeler ise negatif örnek olarak sağlanmıştır. Bunun yanı sıra, raf yüksekliğinin üçte birinden daha küçük yüksekliğe sahip tespitlerin yapılmaması yönünde bir minimum boyut kısıtı tanımlanınca Şekil 4'te görüldüğü gibi yanlış pozitiflerin önemli bir kısmının giderildiği görülmüştür.

Ürün tespiti biriminin çıktısı bize raflardaki ürünlerin konumlarını vermektedir. Bu bilgi kullanılarak ürünlerin



Şekil 4: Ürün tespiti. Minimum eşik değeri tanımlanmadan önce (yukarıda) ve sonra (aşağıda)

etrafındaki sınırlayıcı kutular bulunabilmektedir. Bir sonraki aşama ise bu sınırlar içerisinde hangi markaya ait ürünler bulunduğunu tespit etmektir.

C. Marka Tanıma

Planogram bilgisinde hangi konumda hangi markanın yer alması gerektiği bulunmaktadır. Planogramı bir matris gibi düşünürsek, her hücrenin imgede karşılık gelen bir bölgesi tanımlıdır. Örneğin (2,3) hücresi ikinci raftaki üçüncü ürünü belirtirken, ürün tespiti birimi bu bölgenin bilgilerini (x, y, genişlik, yükseklik) şeklinde verir. Bu bilgilerden yola çıkarak her konumda bulunması gereken markanın orada olup olmadığı sorusu sorulur; yani matrisin her hücresi için doğrulama uygulanır. Dolayısıyla her marka için ayrı bir sınıflandırıcı gerekmektedir. Herhangi bir A markasının sınıflandırıcısında pozitif örnekler A logosu imgeleri, negatif örnekler ise A olmayan diğer markaların logo imgeleridir.



Şekil 5: Bir marka için pozitif örnekler.

Her marka için ayrıca eğitilecek Destek Vektör Makinesi, HOG öznitelikleri ve renk histogramları kullanılarak modellenmiştir. HOG öznitelikleri uzamsal imge piramidinin üç farklı katmanından çıkarılır. Bunun sebebi bütünsel ve yerel özelliklerin ikisini birden yakalayabilmektir. Renk histogramı ise HSV renk uzayından çıkarılır. Renk bilgisinin marka logolarını ayırtmada önemli bir gösterge olduğu varsayılmıştır. Her imgeden çıkarılan bu öznitelikler bir vektörde yan yana sıralanır. Pozitif ve negatif olarak etiketlenmiş logo örneklerinden hesaplanan bu öznitelik vektörleri, Destek Vektör Makinesi için eğitim kümesi olarak kullanılmaktadır. Şekil 5'te bir marka için kullanılan pozitif örnekler gösterilmiştir.

Marka tanıma birimi, daha önceki iki birimin çıktısı olan ızgarayı kullanarak, her bir konum için planograma uygun olup olmadığı ile ilgili bir yanıt üretmektedir.

III. DENEYLER

Bu çalışmada anlatılan birimlerin başarılarını ölçebileceğimiz bir veri kümesi oluşturduk. Öncelikle 4 farklı planograma göre dizilmiş sigara paketlerinin raflar üzerindeki fotoğraflarını flaşlı ve flaşsız çekerek, toplam 229 market imgesi elde ettik. Daha sonra II. bölümde anlattığımız yöntemleri bu veri kümesi üzerinde sınadık. Geliştirmelerde OpenCV görüntü işleme kütüphanesi ve Matlab araç kutularını kullandık.

Ölçümlerin yapılabilmesi için öncelikle veri kümesindeki imgeler için kesin referans (ground truth) oluşturulmuştur. Geliştirilen bir arayüz yardımıyla, bir kullanıcı veri kümesindeki imgeler üzerindeki her raf için iki noktaya tıklayarak doğru denklemi tanımlanması için yeterli olan verileri

sağlamıştır. Daha sonra bu doğrunun orijine uzaklığı ile sistemin tespit ettiği doğrunun orijine uzaklığı arasındaki farka bakılmış ve belirlenen eşikten fazlaysa, raf yanlış tespit edilmiş kabul edilmiştir. Veri kümesindeki 687 raf çizgisinin sistem tarafından tespit edilebilme başarısı iki farklı başarı ölçütü için Tablo I’de verilmiştir. Çizgi bazlı ölçütte, doğru tespit edilen raf çizgisi sayısının toplam raf çizgisi sayısına oranına bakılmıştır. İmge bazlı ölçütte, bütün rafları doğru tespit edilen imge sayısının toplam imge sayısına oranına bakılmıştır. Flaşlı ve flaşsız çekilen fotoğraflarda başarının karşılaştırılması yapılmış ve birinin diğerinden belirgin bir üstünlüğüne rastlanmamıştır.

Planogram	Flaş	Çizgi Bazlı (%)	İmge Bazlı (%)	İmge Sayısı
1	+	91	81	27
	–	94	83	24
2	+	85	69	29
	–	98	93	46
3	+	95	91	22
	–	100	100	23
4	+	99	96	27
	–	69	55	31
Flaşlı		92.38	83.81	105
Flaşsız		90.32	83.06	124
Toplam		91.27	83.41	229

Tablo I: Raf tespiti biriminin başarısı

Ürün tespiti birimini sınamak için veri kümesinden eğitimde kullanılmayan 40 adet raf bölgesi imgesi rastgele seçilmiştir. Bu imgelerdeki toplam 291 sigara kutusunun sistem tarafından tespit edilebilme başarısı ölçülmüştür. Eğitim için 375 pozitif, 198 negatif örnek sağlanmıştır. Pozitif örnekler 20 adet market imgesinden kesilerek işaretlenmiş sigara kutularıdır. Negatif örnekler için ise aynı imgelerden sigara kutusunun tamamını içermeyen bölgeler işaretlenmiştir. Ürün tespiti biriminin başarısı Tablo II’de gösterilmiştir. Buna göre sistem, rafları doğru belirlenmiş imgelerde sigara kutularını %89 kesinlik, ve %88 geri getirme ile bulabilmıştır.

Toplam	Doğru pozitif	Yanlış negatif	Yanlış pozitif
291	256	35	32

Tablo II: Ürün tespiti biriminin 40 raf imgesi üzerindeki başarısı.

Son olarak, marka tanıma birimi sınıflandırma başarısına bakılarak bir sigara markası ile sınanmıştır. Eğitim 80x40 piksel boyutunda 47 pozitif, 87 negatif logo imgesi örneği ile gerçekleştirilmiştir. Sınama ise 31 pozitif, 24 negatif örnek için yapılmış ve %96 sınıflandırma başarısı sonucu alınmıştır.

IV. SONUÇLAR

Bu çalışmada üreticilerin envanter yönetimini kolaylaştırmak ve verimliliği arttırmak amacıyla, market ürünlerinin raflardaki dizilimini otomatik olarak tespit eden bir yöntem sunduk. Bu uygulamada ürünlerin konumlandırılmasının planogram ismi verilen çizenekle olan uyumluluğunu otomatik olarak kontrol etmeyi hedefledik. Önerdiğimiz yapı üç ayrı birimden oluşmaktadır. Hough Dönüşümü ile imge raf sınırları tespit edilerek bölgelere ayrılmış, HOG öznitelikleriyle

Kademeli Nesne Tanıma algoritması kullanılarak ürünlerin yerleri tespit edilmiş, ve bu yerlerde bulunan ürünlerin ait olması gerektiği markaya ait olup olmadığı Desktek Vektör Makinesi ile sorgulanmıştır. Her bir birimin başarısı ayrı ayrı ölçülmüş ve endüstriyel seviyede başarılı sonuçlar elde edilmiştir. Alt birimlerin hepsinin başarı oranları %80’in üzerindedir, bu yönüyle de üreticiler tarafından gerçek hayatta kullanılabilir nitelik taşır.

Gelecekteki çalışmalarda optik karakter tanıma gibi yöntemlerle nesne tanıma güçlendirilerek sistemin geliştirilmesi planlanmaktadır. Ayrıca düzgün yüzeyli biçime sahip olmayan değişik ürün sınıfları (kuruyemiş, cips paketleri vb.) üzerinde çalışmalar yürütülecektir. HOG öznitelikleriyle, daha yavaş fakat daha doğru sonuç verdiği bilinen SIFT özniteliklerinin karşılaştırılması yapılacaktır.

TEŞEKKÜR

Bu bildirideki çalışmalar İdea Teknoloji ve Bilgisayar Çözümleri Ltd. Şti. kurumu tarafından 3130322 nolu TEYDEB projesi kapsamında desteklenmiştir.

KAYNAKÇA

- [1] R. Groenevelt, “Automatic planogram compliance – saving millions of hours of work.” [Online]. Available: <http://mathecsys.com/automatic-planogram-generation/>
- [2] “Planodata: Check planogram compliance through photo recognition,” 2012. [Online]. Available: <http://www.planorama.com/en/planogram-compliance-photo/>
- [3] A. Opalch, A. Fano, F. Linaker, and R. Groenevelt, “Planogram extraction based on image processing,” Patent US 8 189 855, May 29, 2012.
- [4] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [5] J. Kleban, X. Xie, and W. Y. Ma, “Spatial pyramid mining for logo detection in natural scenes,” in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, Jun 2008.
- [6] Y. Akiyama and M. Ito, “Book recognition from color images of book shelves,” in *MVA*, pp. 106–110.
- [7] T. Winlock, E. Christiansen, and S. Belongie, “Toward real-time grocery detection for the visually impaired,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, June 2010, pp. 49–56.
- [8] A. Auclair, L. D. Cohen, and N. Vincent, “How to use SIFT vectors to analyze an image with database templates,” in *Adaptive Multimedia Retrieval*, ser. Lecture Notes in Computer Science, N. Boujemaa, M. Detynecki, and A. Nürnberger, Eds., vol. 4918. Springer, 2007, pp. 224–236.
- [9] R. O. Duda and P. E. Hart, “Use of the Hough transformation to detect lines and curves in pictures,” *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.
- [10] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” 2001, pp. 511–518.