

Koşullu Rastgele Alan Modeli ile Sınır Sahipliği Bilgisi Çıkarımı

Extraction of Border Ownership Information by Conditional Random Field Model

Buğra Özkan
Bilgisayar Mühendisliği Bölümü
Orta Doğu Teknik Üniversitesi
Ankara, Türkiye
bozkan@ceng.metu.edu.tr

Sinan Kalkan
Bilgisayar Mühendisliği Bölümü
Orta Doğu Teknik Üniversitesi
Ankara, Türkiye
skalkan@ceng.metu.edu.tr

Özetçe— Sınır sahipliği bilgisi, görüntü üzerindeki kenarların hangi görüntü alanlarına ait olduğunu belirleyen bir bilgi olup, son zamanlarda, bilgisayarlı görme alanında da kullanılmaya başlanmıştır. Figür-zemin ayrımı, derinlik tahmini ve optik akış gibi bilgisayarlı görme alanlarında kullanılmaya başlanan bu bilgi, insan görme sistemlerinin analizi sonucu çıkarılan modeller doğrultusunda oluşturulmakta olup, literatür henüz yeterli seviyede modeller üretememiştir. Bu çalışmada amaç, sınır sahipliği bilgisini, komşu görüntü alanları arasındaki içeriksel ilişkileri Koşullu Rastgele Alan ile modelleyerek, başlangıçta bazı şekilsel öznitelikler doğrultusunda etiketlenmiş sınırların etiketlerinin bu model yardımıyla iyileştirilmesini sağlamaktır.

Anahtar Kelimeler — sınır sahipliği, figür-zemin ayrımı, koşullu rastgele alan, grafiksel modeller

Abstract— Border ownership is a kind of information used for determining the regions that own borders in an image. This information has recently become more valuable as it began to be used in many vision problems such as figure-ground segregation, depth perception and optical flow, however the quality and quantity of current literature are not sufficient yet. In this study, a Conditional Random Field model is developed for the enhancement of border ownership labelings, which are extracted with the help of some spatial cues at the beginning.

Keywords —border ownership, figure-ground segregation, conditional random field, graphical models

I. GİRİŞ

İnsan görme sistemi, üç boyutlu gerçek dünya sahnesini iki boyutlu görüntü olarak algılayarak bilgi kaybına neden olur. Boyut değişiminden kaynaklanan derinlik bilgisi gibi kaybolan görsel bilgiler, görme sistemimiz tarafından düzeltilmektedir. Tekdüze koyuluğa veya dokuya sahip, ayırt edici bir bilgiye sahip olmayan bir görüntü alanı, yine de algılanabilmektedir. Eksik olan bu bilginin düzeltilmesi, ilgili görüntü alanlarını çevreleyen sınırlardaki görsel bilgilerin 'içeriye doldurma' mekanizması sayesinde bu alanlardaki bilgilerin iç kısımlara iletilmesiyle mümkün hale gelmektedir [1]. Yapay görme sistemleri alanındaki ilgili çalışmalar da, aynı problemi, insan görme sistemlerinin bu yapısını modelleyerek çözmeye

çalışmaktadırlar. Fakat bu modeller, hangi sınırların hangi görüntü alanlarına ait olduğunu bilmemizi gerektirmekte olduğundan, 'sınır sahipliği' olarak adlandırılan görsel bir bilgi türünü ortaya çıkarmaktadır.

Sınır sahipliği, bir görüntüdeki kenarların hangi görüntü alanlarına ait olduğunu ifade eden bir bilgidir. Bu bilginin çıkarımı, fizyolojik ve psikolojik deneyler sonucu elde edilen bilgiler üzerine kurulan modeller ile gerçekleştirilmektedir. Bu öznitelikler, şekil ve renk tabanlı olmak üzere iki ana sınıf altında değerlendirilebilir. Başlıca şekil tabanlı öznitelikler, alan sınırının yönelim, dışbükeylik gibi özellikleri ile bu sınıra komşu alanların simetrisi ve boyutu gibi etkenlerdir[3]. Ek olarak, L ve T köşeleri gibi şekilsel yapılar da sınır sahipliği bilgisinin çıkarılmasında önemli role sahiptir. Her ne kadar, Nishimura ve Sakai, T-köşeleri gibi görsel yapıları hassasiyeti bulunan beyin hücrelerinin varlığı henüz bilinmediğinden, sınır sahipliği modelleri için bu yapıların kullanılmaması gerektiğini beyan etse de [4], bu alandaki son çalışmalardan birisinde Leichter ve Lindenbaum [3], eğrilik özniteliği ile beraber kullanılan T-köşeleri bilgisinin model performansını önemli bir ölçüde arttırdığını deneyleri sonucu göstermişlerdir.

Sınır sahipliği görece yeni bir çalışma konusu olduğundan, bu alandaki çalışmalar yeterli sayıda değildir. Var olan çalışmalar ise genellikle hiç bir eğitim verisi ve algoritması kullanılmadan, yerel kısıtlar üzerine tanımlanmış deneysel çalışmalar olmakla beraber, basit sentetik veriler üzerinde test edilmiştir. Bu çalışmalar, sonuçlarını birkaç temel şeklin bulunduğu resimler üzerinde göstererek savundukları teorileri ispatlamayı yeterli görmektedir. Gerçek görüntü üzerinde çalışan ve Koşullu Rastgele Alan ile sınır sahipliği modeli oluşturan Leichter ve Lindenbaum [3], başlangıç etiketlerini, oluşturdıkları 2.1D model doğrultusunda, görüntü alanlarının sıralı derinlik bilgileri ile atamakta olup, test verisi olarak dış mekan resimlerini seçmiştir. Ancak derinlik bilgisi, dış mekana kıyasla iç mekanda algılanması daha zor bir bilgi olduğundan, iç mekan görüntüleri için bu çalışmanın aynı başarıyı yakalaması beklenemez. Bu çalışma ise, iç ve dış mekan fark etmeksizin, ortak sınıra sahip iki görüntü alanı arasındaki içeriksel ilişkileri başarılı bir şekilde modelleyip, bu grafiksel model sayesinde

başlangıç etiketlerini mümkün olduğunca iyileştirmeyi amaçlamaktadır.

II. VERİ KÜMESİ

Algoritmanın eğitim ve testinde kullanılan görüntüler, LHI ve Berkeley veri kümelerinden edinilmiştir [5,8]. Eğitim için 100, test için 20 farklı görüntü kullanılmıştır. Kullanılan bu 120 resimlik veri kümesi, hem iç, hem dış mekan görüntülerinden oluşmakta olup, görüntü alanlarının el ile çizilerek işaretlenmiş olduğu doğruluk referans maskelerini de içermektedir. Eğitim veri kümesindeki görüntü sınırları el ile ait oldukları görüntü alanlarına atanarak sınır sahipliği bilgisi oluşturulmuştur. Veri kümesinden örnek bir resim ile görüntü alanlarının çıkarılmış olduğu maske sırasıyla Şekil 1'de görülebilir. Görüntü alanlarının ayrı ayrı renklendirilmiş olduğu maskede koltuk (*mavi*) ve pencere (*eflatun*) arasındaki sınır beyaz ile gösterilmiş olup, bu örnekte sınır, önde olan koltuğa aittir.



Şekil 1: Örnek resim ve ilgili görüntü alan maskesi

III. YÖNTEM VE ALGORİTMA

Sınır sahipliği bulma algoritmamız temel olarak, T ve L köşeleri yardımıyla sınır sahipliği bilgisi çıkarılan test verisinin sınır sahipliği istatistikî olasılığının, sınırları hazır olarak sahibi olan alanlara atanmış eğitim veri kümesi ile oluşturulan Koşullu Rastgele Alan (KRA) modeli sayesinde artırılmasına dayanır. Bu model, kontrast oranı ve entropi, T ve L köşeleri ile alt alan olmak üzere iki spektral ve üç şekilsel ipucu üzerine kurulmuştur.

Çalışmada, görüntü alanları sınır sahipliği bilgisinin olasılık dağılım fonksiyonu bir Gibbs dağılımı şeklinde modellenmekte olup, bu ikili KRA modelini tekli ve ikili klik potansiyellerinin toplamları oluşturmaktadır. Bu potansiyel fonksiyonları, grafik üzerindeki tek bir düğümün veya beraber

komşu iki düğümün veri-etiket ilişkisini nasıl etkilediklerini gösterir. Potansiyel parametrelerinin tahmini sonrasında bu model yardımıyla ardıl olasılıkların hesaplanması mümkün hale gelir. Sınır sahipliği için geliştirilen KRA modeli şu şekilde formüle edilebilir:

$$P(R|x, \sigma) \rightarrow - \sum V1(Ri, x, \sigma) - \sum V2(Ri, Rj, x, \sigma) \quad (1)$$

Her ikili KRA modelinde olduğu gibi modelimizde, veriye (x) ait sonsal sınır sahipliği olasılığının değeri (R) hesaplanırken, tüm tekli (V_1) ve ikili (V_2) potansiyellerin toplamı bulunur ve bu toplamın negatif değerinin üsteli normalize edilerek hesaplanır. Potansiyeller, tek bir görüntü alanı ile iki komşu görüntü alanının verinin etiketlenmesinde nasıl rol aldığını, σ parametreleri ile belirler.

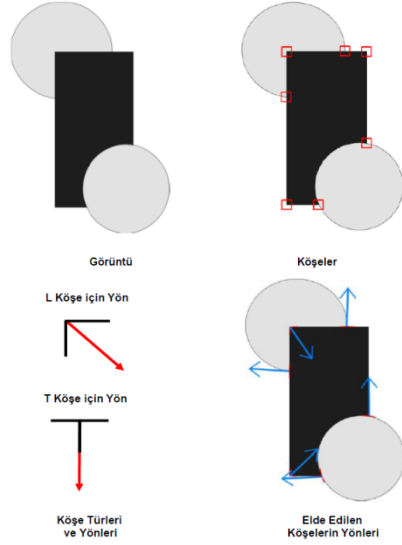
Çalışmada kullanılan sınır sahipliği modeli, iki düğümünden (node) oluşan sade bir grafiktir. Bu modelin düğümleri, ortak bir sınıra sahip iki komşu görüntü alanını tanımlar. Sınır sahipliği ilişkisi komşu iki alan üzerinden doğan bir ilişki olduğundan dolayı, bu basit grafik, sınır sahipliği modeli için yeterlidir. Her iki komşu görüntü alanı, içeriksel ilişkileri doğrultusunda aralarında bulunan sınır için sahiplik değerini beraber belirler, bu sebeple bir KRA modeli, sınır sahipliği problemini çözmek için en uygun yöntem olarak gözükmektedir. Bu içeriksel ilişkiler için en ayırt edici öznelikler kullanmak, daha iyi bir model kurmamızı sağlar. Grafik düğüm ve kenarları için kullanılan bu öznelikler, model parametrelerinin hesaplanması amacıyla tekli ve ikili potansiyellere dönüştürülür.

Sınır sahipliği bulma algoritmamız üç ana basamaktan oluşmaktadır. Bunlar sırasıyla:

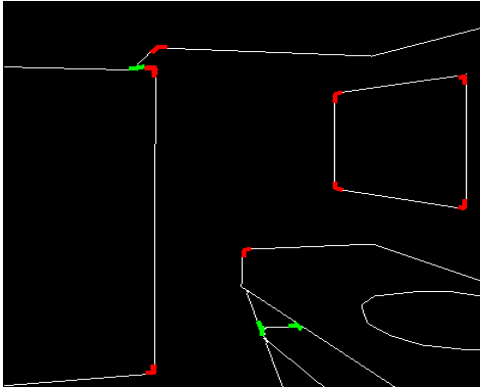
1. Test ve eğitim veri kümesi görüntülerindeki tüm düğüm ve kenar özneliklerinin çıkarılması, (*Bu özneliklerden T ve L köşeleri, test resmi için ilk etiketlendirme amacıyla da kullanılacaktır.*)
2. Eğitim veri kümesi ile KRA modelinin oluşturulması,
3. Test görüntüsünden çıkarılan sınır sahipliği önsel olasılıklarının KRA modeli ile iyileştirilmesidir.

Algoritmanın ilk adımında, sınırları işaretlenmiş ve etiketlenmiş test görüntüsünde sınırlar üzerindeki T ve L köşeleri çıkarılır. Bulunan T ve L köşeleri yönleri, üzerinde buldukları sınırın hangi görüntü alanına ait olduğunu belirlemede kullanılmaktadır. Örnek bir görüntü üzerinde bu köşelerin sınırlar üzerinde nerelere denk geldiği, türleri ve yönleri ile Şekil 3 üzerinde detaylı olarak gösterilmiştir. L köşesi yönündeki alanlar ile T köşesi zıt yönündeki alanlar bize çoğunlukla sınır sahibi alanı vermektedir. L köşesi, 2 kenar kesişiminden oluşmakta olup, aralarındaki açı yaklaşık olarak en fazla 100-110 derece olmalıdır. T köşesi ise, birbirine dik 2 kenardan oluşup 3 ayrı bölge arasında sınır oluşturmaktadır. Bu çalışmada yapılan veri analizleri ve ilgili literatürdeki çalışmalar [3,6] özellikle T köşesinin performans değerlerini önemli ölçüde arttırdığını göstermektedir. Görüntü alan sınırları hazır olarak mevcut olduğundan, bu sınırlar üzerindeki T ve L köşelerini bulmak için basit morfolojik

operasyonlar yeterli olmuştur. Görüntü alan sınırlarının işaretlenmiş olduğu örnek bir veri üzerindeki T ve L köşeleri, Şekil 4'te gösterilmiştir. Bu örnek resim üzerinde T köşeleri yeşil, L köşeleri ise kırmızı renkte işaretlenmiştir.



Şekil 3: T ve L köşeleri hakkında görsel bilgi



Şekil 4: T ve L köşeleri örnek sonucu

T köşelerinin L köşelerine oranla daha ayırt edici bir öznitelik olmasından dolayı, test ve eğitim verisindeki sınırlara görüntü alanı aidiyet değeri atanırken öncelikle T köşeleri göz önünde bulundurulmakta olup, T köşelerinden aidiyet değeri atanamaması durumunda L köşeleri üzerinden çıkarım yapılmaktadır. Köşe sayısı ile önsel olasılık çıkarımı şu şekilde formülize edilebilir:

$$A(i, j) = \text{if } \sum \text{Dir}T(i, j) \neq 0 \\ \text{then } \text{sign}(\text{Dir}T(i, j)) \\ \text{else } \text{sign}(\text{Dir}L(i, j)) \quad (1)$$

İki komşu alan, B_i ve B_j arasındaki sınırdaki bulunan bir T veya L köşesinin yönü ($\text{Dir}T$ & $\text{Dir}L$) B_i alanını gösteriyorsa +1, B_j alanını gösteriyorsa -1 değerini alır (1). Sınır üzerindeki tüm T köşelerinin gösterdiği ortak yön, T köşesi yön

değerlerinin toplamının pozitif veya negatif olması ile belirlenir ($\text{sign}(\text{Dir})$). T köşelerinden bir yön elde edilememesi durumunda ($\sum \text{Dir}T(i, j) \neq 0$) ise tüm L köşelerinin gösterdiği yön, sınır-alan aitlik değerini belirler. Örneğin, Şekil 4'te resmin sağ üst tarafında bulunan dörtgen alan etrafındaki sınırın önsel aitlik değeri, yani bu sınırın dörtgen içerisindeki alana mı, dışarısındaki alana mı ait olduğunu belirlerken, öncelikle sınır üzerindeki T köşelerine bakılır. Sınır üzerinde bu tür köşe mevcut olmadığından, L köşelerinin gösterdiği yön (iç alan), sınır sahibi alan olarak adlandırılır.

Köşe sayısına göre önsel olasılıkları atanan KRA modelinin eğitimi için düğüm ve kenar öznitelikleri çıkarılır. Düğüm öznitelikleri ve özniteliklerin ayırt edici yönleri şunlardır:

- **Kontrast oranı:** Sınır etrafındaki görüntü alanı parçasının kontrast ortalama değerinin tüm görüntü alanının kontrast ortalama değerine oranıdır. Canlılar üzerinde yapılan çalışmalar [7], V2 görme alanının sınırı etrafındaki kontrast farklılığını sınır sahipliği için önemli bir ipucu olarak göstermiştir.
- **Entropi:** Görüntü alanı ortalama entropi değeridir. Çalışma süresince yapılan deneysel çalışmalar göstermiştir ki, arka plandaki görüntü alanları öndeki alanlara göre genelde daha az doku ve rastgelelik içermektedir.
- **Paralel T köşe sayısı:** Görüntü alanı sınırları üzerinde bulunup, geniş kenarı sınır ile paralel olan toplam T köşe sayısıdır. Bu doğrultudaki T köşeleri [3], sınır sahipliği konusunda önemli ölçüde başarı sağlamaktadır.
- **L köşe sayısı:** Alan sınırları üzerinde bulunan toplam L köşe sayısıdır. L köşeleri, genellikle önde bulunan, yani sınırın sahibi olan görüntü alanları üzerinde bulunmakta olup, sınır sahibi alanların ayırt edici bir özelliğidir.
- **Alt alan özelliği:** Görüntü alanının görüntü üzerindeki pozisyonu, sınır sahipliği için bir diğer önemli şekilsel ipucudur. Üstte kalan, yani sınır sahibi olan alan, genellikle diğerine oranla, görüntüde daha üstte bulunur. Bu özellik, çalışmamızda, görüntü alanının en düşük dikey koordinatının resmin boyuna oranı olarak alınarak normalize edilmiştir.

Kenar özniteliği olarak ise, çoğu KRA modelinde de kullanıldığı üzere, kenarın iki ucundaki düğüm öznitelikleri kullanılmıştır. Model parametre tahmininden sonra, değişkenlerin ortak olasılığının en yüksek değerini bulabilmek için düğüm çözme(decoding) işlemi uygulanır. Test aşaması olan düğüm çözme işlemi sonucunda görüntü alanlarının yeni sınır sahipliği bilgileri edinilmektedir.

IV. DENEYSEL SONUÇLAR

Algoritma testi sonucunda elde edilen performans değerleri Tablo 1'de görülebilir. Bu değerler, 20 test resminin içerdiği toplam 335 adet görüntü sınırına ait yüzdesel doğruluk sonucunu göstermektedir. Tablodaki yüzde kısmında ise,

doğru işaretlenmiş sınır sayısının resim üzerindeki tüm sınır sayısına oranı belirtilmiştir. Örnek bir resim üzerinde algoritma sonucu ise Şekil 5'te görülebilir. Bu resim üzerinde sınır sahipliği doğru görüntü alanına atanan sınırlar yeşil, yanlış atananlar ise kırmızı olarak renklendirilmiştir.

Test Görüntüsü	Sonuçlar		
	Toplam sınır sayısı	Doğru etiketlenen sınır sayısı	Yüzde
İç alan	132	86	65,1
Dış alan	203	146	71,9
Toplam	335	232	69,3

Tablo 1: Sonuç tablosu



Şekil 5: Örnek algoritma çıktı resmi

V. SONUÇ

Bu çalışmada, komşu görüntü alanları arasındaki sınırların hangi alana ait olduğunu bulma problemi olan sınır sahipliği problemi üzerinde çalışılmış olup, ortak sınıra sahip alanların birbiriyle ve tüm resim verisiyle ilişkilerini modellemek için istatistiksel bir modelleme yöntemi olan Koşullu Rastgele Alan'ın bu problem için en uygun aday olduğu sonucuna varılmış olup, deneylerde elde edilen sonuçlar da bu tezi doğrulamıştır. Renksel ve şekilsel bir takım öznitelikler kullanılarak komşu görüntü alanları arasındaki içeriksel ilişkiler modellenmeye çalışılmış olup bu model, Berkeley ve LHI görüntü kümelerinden toplanan 20 farklı iç ve dış mekan resmi üzerinde denenmiş ve %69'luk bir başarı yüzdesiyle sonuç vermiştir. Sınır sahipliği çalışmalarının büyük bir çoğunluğu, gerçek görüntüler üzerinde test edilmediğinden makul bir karşılaştırma yapmak söz konusu değildir. Leichter ve Lindenbaum [3] ise, gerçek dış mekan görüntüleri üzerinde %82'lik bir başarı oranı elde etmiş olmalarına rağmen, modellerini obje derinlik bilgisi üzerine kurmaları ve bu bilginin iç mekan görüntülerinden sağlıklı elde edilemeyeceği sebebiyle, her tür görüntüde aynı başarının elde edilemeyeceği öngörülmektedir. Bu çalışmada ise asıl amaçlanan, sınır sahipliği bilgisinin mekan gibi etkenlere bağlı kalmadan elde edilebilmesidir. Nitekim Tablo 1'deki sonuçlar da bu hedefin gerçekleştirildiğini gösterir niteliktedir.

Sınır sahipliği algoritmasında, kontrast ve entropi oranı gibi öznitelikler için seçilen parça alan deneysel olarak belirli bir oran büyüklüğünde seçilmekte olup, bu oran mantıklı bir değer dönmemektedir. İlerideki çalışmalarımızda bu gibi sezgisel değerlerin de öğrenilmesi, yapılacak ilk işlerden olacaktır. Yeni öznitelikler modele katılacak, tüm bu özniteliklerin modele nasıl ve ne oranda katkı sağladığı incelenecektir. Şekil 5'teki örnek sonuç resminde görüldüğü üzere, kürek ve göl arasındaki sınır göle atandığından yanlış bir etiketlendirme söz konusudur. Bu örnekte kürek, kontrast ve dokusallık öznitelikleri için arka plan özelliği yansıttığından sınır yanlış etiketlenmiştir. Bu gibi örneklerde şekilsel özniteliklere daha çok ağırlık verilmelidir. Resim içeriği doğrultusunda özniteliklere ağırlık atama, bu problemi çözmek için bir çözüm olabilir. Ek olarak, sınırlar için yapılan ilk etiketlendirme, şekil tabanlı sayısal bir limit üzerine olup, bu sınıflandırmanın da Destek Vektör Makineleri gibi ileri seviye bir öğrenme yöntemiyle yapılması düşünülmektedir.

BİLGİLENDİRME

Bu çalışma Tübitak tarafından desteklenen 111E155 numaralı proje kapsamında yapılmıştır.

KAYNAKÇA

- [1] Pessoa, L., Thompson, E., & Noë, A., "Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception", *Behavioral and Brain Sciences*, 21(6), 723-748, 1998.
- [2] Kikuchi, M., & Akashi, Y., "A model of border-ownership coding in early vision", *Artificial Neural Networks—ICANN*, 1069-1074, 2001.
- [3] Leichter, I., & Lindenbaum, M., "Boundary ownership by lifting to 2.1d", In *IEEE 12th International Conference on Computer Vision*, (pp. 9-16). IEEE, 2009.
- [4] Nishimura, H., & Sakai, K., "Determination of border ownership based on the surround context of contrast", *Neurocomputing*, 58, 843-848, 2004.
- [5] Yao, B., Yang, X., & Zhu, S. C., "Introduction to a large-scale general purpose ground truth database: methodology, annotation tool and benchmarks", In *Energy Minimization Methods in Computer Vision and Pattern Recognition* (pp. 169-183). Springer Berlin/Heidelberg, 2007.
- [6] Kalkan, S., Yan, S., Pilz, F., & Krüger, N., "Improving junction detection by semantic interpretation", *VISAPP (1)*, 264-271, 2007.
- [7] Zhou, H., Friedman, H., von der Heydt, R., "Coding of Border Ownership in Monkey Visual Cortex", *The Journal of Neuroscience*, 20(17):6594-6611, September 1, 2000.
- [8] Martin, D., Fowlkes, C., Tal, D., & Malik, J., "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics", In *Proceedings. Eighth IEEE International Conference on Computer Vision, ICCV (Vol. 2, pp. 416-423)*. IEEE, 2001.