

Kinect RGB Görüntülerde ve Derinlik Haritalarında Uzam-zamansal Özellikleri Kullanarak Türk İşaret Dili Tanıma

Turkish Sign Language Recognition Using Spatio-temporal Features on Kinect RGB Video Sequences and Depth Maps

Abbas Memiş
Bilgisayar Mühendisliği Bölümü
Yıldız Teknik Üniversitesi
İstanbul, Türkiye
abbasmemis@gmail.com

Songül Albayrak
Bilgisayar Mühendisliği Bölümü
Yıldız Teknik Üniversitesi
İstanbul, Türkiye
songul@ce.yildiz.edu.tr

Özetçe—Bu bildiri, Kinect duyargaya ait RGB video görüntülerde ve derinlik haritalarında uzam-zamansal özellikleri kullanan bir Türk İşaret Dili tanıma sistemini sunmaktadır. Sistemde, ardışıl video görüntülerdeki dinamik işaretlerin zamansal özelliklerini ifade etmek için, hareket farklarına dayalı toplamsal imge yaklaşımı kullanılmaktadır. Toplamsal imgeler, işaret süresince işaretçinin gerçekleştirdiği hareketin bütününi ifade etmektedir. Hareketlere ait uzamsal özellikleri elde etmek için, toplamsal imgelere 2-B Ayrık Kosinüs Dönüşümü (DCT) uygulanmakta ve işaretlerin enerji yoğunluklarını gösteren dönüşüm imgeleri elde edilmektedir. Yöntemler, işaretlere ait RGB video görüntülere ve derinlik haritalarına ayrı ayrı uygulanıp iki adet dönüşüm imgesi elde edilmektedir. İşaretlere ait öznitelik vektörlerinin oluşturulması için ise dönüşüm imgelerinde zig-zag tarama ile yüksek enerjili DCT katsayıları belirli oranlarda seçilerek birleştirilmektedir. Sistemin tanıma aşamasında, Manhattan uzaklığını kullanan K-en yakın komşu sınıflandırıcısı kullanılmıştır. Sistem performansının değerlendirilmesi amacıyla Türk İşaret Dili'ne (TİD) ait 3 ayrı kategoride toplam 111 kelimededen oluşan 1002 işaretlik bir veri seti kullanılmıştır. Çalışma sonucunda yaklaşık %90 oranında bir tanıma başarımu gözlemlenmiştir.

Anahtar Kelimeler—*işaret dili tanıma; Türk İşaret Dili; dinamik işaretler; uzam-zamansal özellikler; Kinect duyarga; derinlik haritaları.*

Abstract—This paper presents a Turkish Sign Language recognition system that uses spatio-temporal features on Kinect sensor RGB video sequences and depth maps. Proposed system uses cumulative motion images which based on motion differences and represent the temporal characteristics of dynamic signs in motion sequences. Cumulative motion images represent the whole motions of signers. 2-D Discrete Cosine Transform (DCT) is applied to cumulative sign images in order to obtain spatial features of signs and transformed images that represent the energy density of signs are obtained. Two transform images are obtained by applying referred methods to both of RGB video sequences and depth maps separately. Feature vectors of dynamic

signs are produced by combining a certain amount of DCT coefficients that contain higher energy via zig-zag scanning on transform images. K-Nearist Neighbor classifier with Manhattan distance used for recognition process. System performance is evaluated on a sign database that contains 1002 signs belongs to 111 words in three different categories of Turkish Sign Language (TID). Proposed sign language recognition system has a recognition rate about %90.

Keywords—*sign language recognition; Turkish Sign Language; dynamic signs; spatio-temporal features, Kinect sensor; depth maps.*

I. GİRİŞ

İşaret dilleri işitme ve konuşma engelliler tarafından kullanılan; ellerin, kolların hareket ve şekilleri ile yüz ifadelerinden oluşan görsel bir dildir. İşaret dillerine ait kelimeler daha çok ellerin kullanıldığı, ellerin zaman içerisinde şekil ve yer değiştirmesi ile meydana gelen sürekli işaretlerden oluşur. Engelliler için bir anadil olmasıyla birlikte, konuşma dilleri gibi, ülkelere göre değişiklikler gösterir [1].

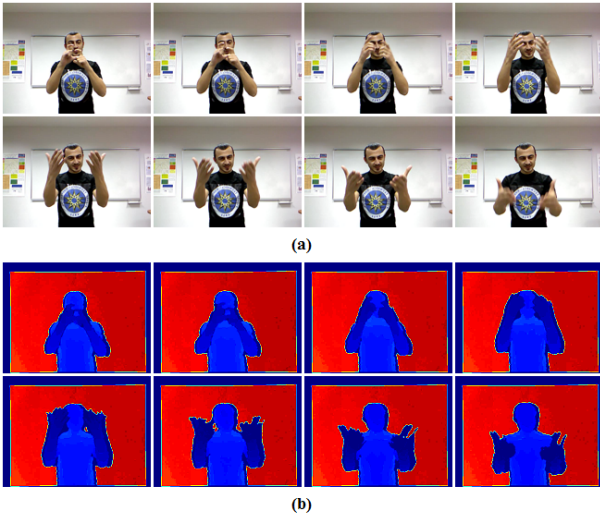
İşaret dillerini tanıma amacıyla yapılan çalışmalarda, elektronik algılayıcıları eldiven ve gözü tabanlı farklı yaklaşımlar mevcuttur. Öz ve Leu [2] sensörlü eldivenler kullanarak Amerikan İşaret Dili'ne ait 50 kelime üzerinde yaptıkları çalışmada, yapay sinir ağları yaklaşımıyla kelimeleri yaklaşık %90 başarımla tanımışlardır. Gao, Fang, Zhao ve Chen [3] Çin İşaret Dili üzerine yaptıkları öz düzenleyici özellik haritaları ve HMM yaklaşımıyla, yine algılayıcı eldiven verilerini kullanarak 5113 kelimedede yaklaşık %87 tanıma oranına ulaşmışlardır.

Görü tabanlı yaklaşımlarda, ellerin konum ve şekillerinin zamanda önceki ve sonraki durumlarına bağlılığı sebebi ile Markov modellerin kullanımı yaygındır. Starner tarafından Kinect duyarga kullanılarak yapılan bir çalışmada [4], HMM

yaklaşımıyla Amerikan İşaret Dili'ne ait 494 deyim %75 civarında başarımla tanınmıştır. Haberdar ve Albayrak [5], HMM ile genel ve yerel özellikleri kullanarak 50 kelime üzerinde yaptıkları çalışmada %95.4 tanıma gözlemlenmiştir. Shanableh [6] Arap İşaret Dili'ne ait hareketleri tanımak amacıyla hareket kestirimleri ile uzam-zamansal özellikleri farklı sınıflandırıcı modellerinde kullanarak 23 kelimelik bir veri setinde %97 ile %100 arasında değişen başarılı sonuçlar elde etmiştir.

İşaret dillerindeki kelimelerin ifadesinde kullanılan yöntemlerden biri de parmak hecelemedir. Türk İşaret Dili'nde parmak heceleme tanıma amacıyla yapılan çalışmalar [7] mevcuttur. Altun ve Albayrak tarafından [7]'de önerilen sistemde genellenmiş Hough dönüşümü, ilgi bölgeleri ve yerel tanımlayıcılar yaklaşımı ile 29 durağan işaret %93 oranında tanınabilmiştir.

Bildiride sunulan çalışmada, Kinect duyargaya ait RGB video görüntülerde ve derinlik haritalarında Türk İşaret Dili'ne ait dinamik işaretleri tanımak için uzam-zamansal özellikler kullanılmıştır. Bildirinin II. bölümünde ön-işleme adımları belirtilmiştir. Özellik çıkarımı aşamasında yapılan işlemler bölüm III'te detaylandırılmıştır. IV. bölümde deneysel çalışmalara, V. bölümde ise sonuçlara yer verilmiştir.



Şekil 1. Türk İşaret Dili'nde “günaydın” kelimesinin ilgili veri setinde bir işaretçi tarafından ifade edilmesine ait video görüntü (a) ve derinlik haritası (b) örnek çerçeveleri (Derinlik haritaları renklendirilmiş olarak).

II. GÖRÜNTÜ ÖN-İŞLEME

Çalışmada kullanılan dinamik işaretlere ait video görüntüler üzerinde ilk ön-işleme adımı olarak çerçeve boyutu küçültme işlemleri yapılmıştır. Gerçek boyutları 640×480 olan görüntü çerçeveleri ve derinlik haritaları, $1/2$ ve $1/4$ ölçeklerinde yeniden boyutlandırılarak sırasıyla 320×240 ve 160×120 boyutlarında görüntüler elde edilmiştir. Böylece farklı çerçeve boyutlarında tanıma performansları gözlemlenmiştir. Bu işlemi takiben görüntülerin bulunduğu renk uzayı değiştirilmiştir. Görüntüler RGB renk uzayından gri-ölçek uzaya aktarılmıştır. Bu adımlardan sonra özneliklerin çıkarılması sürecindeki işlemlerle devam edilmiştir. Şekil 1'de çalışma kapsamında değerlendirilen “günaydın” kelimesinin RGB görüntülerine ve derinlik haritalarına ait örnek çerçeveler gösterilmiştir.

III. ÖZNETELİK ÇIKARIMI

Çalışmada kullanılan dinamik işaretlere ait özneliklerin çıkarımı amacıyla, imge ön işleme adımlarını takiben, işaretlerin video görüntülerini oluşturan ardışıl çerçevelerdeki hareket farklılıkları değerlendirilmiştir. İşaretlere karşılık gelen RGB görüntü ve derinlik haritalarının ardışıl çerçevelerden elde edilen hareket değişimleri toplanarak her bir işaret için bütün işareti temsil eden iki adet toplamsal hareket imgesi oluşturulmuştur. Elde edilen toplamsal imgelerdeki küçük hareket değişimlerini süzgeçlemek amacıyla medyan filtresi kullanılmıştır. Harekete ait zamansal özellikleri içeren toplamsal imgeden uzamsal özellikleri elde etmek amacıyla, 2-boyutlu (2-D) Ayrık Kosinüs Dönüşümü uygulanmıştır. Elde edilen dönüşüm imgelerinden, yüksek enerji içeren DCT katsayıları zig-zag tarama ile değişik oranlarda seçilerek işaretlere karşılık düşen öznelik vektörleri oluşturulmuştur.

A. Hareket Farkları ile Zamansal Özneliklerin Çıkarımı

Video görüntülerdeki ardışıl çerçevelerin hareket farklarının çıkarımı video sıkıştırmada kullanılan ara yöntemlerden biridir. Yapılan çalışmada, bu yöntem işaret diline ait sürekli işaretlerin zamana bağlı özelliklerini çıkarmak ve ifade etmek amacıyla kullanılmıştır. Bu işlem işaretlere ait görüntülerdeki tüm ardışıl çerçevelerdeki parlaklık değerlerinin farkları alınarak (1) gerçekleştirilmiştir.

$$I_{F_i} = I_{i+1} - I_i \quad (i = 1, \dots, n-1) \quad (1)$$

(1)'deki denklemde I_i ve I_{F_i} sırasıyla işarete ait i . video görüntü çerçevesini, ve ardışıl iki çerçeve arasındaki fark imgesini ifade etmektedir. İşlem sonucunda elde edilen fark imgelerinin her biri ikili-eşikleme ile ikili imgelere dönüştürülmüş ve ikili imgelerin toplamı alınarak işaretin zamansal özelliklerini ihtiva eden bir toplamsal imge oluşturulmuştur. (2)'de gösterilen denklemde, I_{E_i} eşikleme uygulanmış fark imgelerini, T eşik katsayısını, n toplam çerçeve sayısını ve I_T ise toplamsal imgeyi temsil etmektedir.

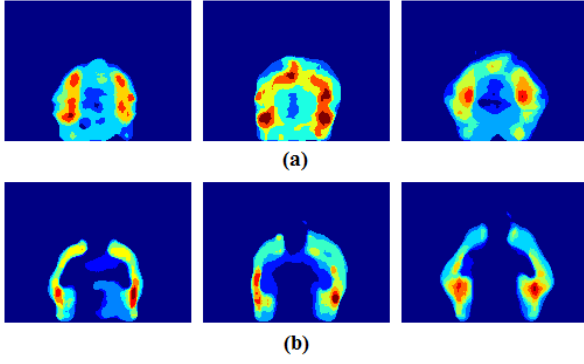
$$I_{E_i} = \begin{cases} 1 & \text{eğer } |I_{F_i}| \geq T \\ 0 & \text{değilse} \end{cases} \quad I_T = \frac{1}{n} \cdot \sum_{i=1}^{n-1} I_{E_i} \quad (2)$$

Elde edilen toplamsal imgelere medyan filtresi uygulanarak, hareket bölgelerindeki durağan gibi görünen küçük bölgeler giderilmiştir. İşlemler sonucunda sürekli işaretlerin zamandaki özelliklerini içeren iki-boyutlu toplamsal imgeler elde edilmiştir. Şekil 2'de çalışma kapsamında 3 farklı işaretçi tarafından gerçekleştirilen “iyi akşamlar” kelimesine ait örnek gri-imge ve derinlik haritası toplamsal imgeleri verilmiştir.

Şekil 2'de yer alan imgelere sıcak renkler (kırmızı tonları) hareketlerin yoğunlaştığı, soğuk renkler (mavi tonları) ise hareketin olmadığı veya az olduğu bölgeleri göstermektedir.

B. Uzamsal Özneliklerin Çıkarımı ve Öznelik Vektörlerinin Oluşturulması

İşaretlere ait zamansal özelliklerin çıkarılmasını takiben elde edilen video ve derinlik haritası toplamsal imgelerine 2-boyutlu DCT dönüşümü uygulanmıştır. DCT ile dinamik



Şekil 2. 3 farklı işaretçi için “iyi akşamlar” kelimesinin gri-imge farklarından elde edilen toplamsal imge (a) ve derinlik haritası farklarından elde edilen toplamsal imge (b) (0-255 aralığına normalleştirme uygulandı ve renklendirildiği durumda).

işaretlerin zamansal özellikleri uzamsal uzaya aktararak, uzamsal özellikler elde edilmiştir. DCT işlemi sonucunda elde edilen dönüşüm imgeleri, öznitelikleri elde etme aşamasında kullanılmaktadır.

DCT, karakteristiği gereği, uygulandığı imgelerin yüksek enerjisini, katsayılar matrisinin sol-üst köşesinde toplar ve imgeyi daha küçük boyutta veri ile temsil etmeyi sağlar. Bu özelliği ile veri sıkıştırma da kullanılmaktadır. Yapılan çalışmada, toplamsal imgelere DCT uygulanarak öznitelik vektörleri elde edilmiştir. DCT katsayıları, gri-imge ve derinlik haritaları dönüşüm imgelerinin sol-üst köşesinden başlanarak zig-zag tarama ile okunup belirli oranlarda birleştirilerek, işaretlerin tanınması aşamasında kullanılacak olan asıl öznitelik vektörleri oluşturulmaktadır.

IV. DENEYSEL ÇALIŞMALAR

Yapılan çalışmada dinamik işaretleri tanımak için Türk İşaret Dili'ne (TİD) ait 111 kelimededen oluşan bir veri seti oluşturulmuştur. Veri setinde yer alan kelimeler; günlük konuşma, zaman ve meslek terimlerini içeren üç kategoride toplanmış ve kelimeler işitme engelliler tarafından yapılan günlük konuşmada ihtiyaç duyulan kelimelerden oluşmaktadır. İşaretler Kinect duyarga kullanılarak kayda alınmış ve işaretlerin video görüntüleri ile birlikte derinlik haritaları da elde edilmiştir. Veri setindeki her kelimenin; 3 farklı işaretçinin 3 tekrarından oluşan 9 adet kaydı bulunmaktadır. İşaretlere ait bazı örneklerin değerlendirme dışı bırakılmasıyla toplamda 1002 adet örnek değerlendirilmeye dahil edilmiştir. Tablo I'de çalışma kapsamında kullanılan farklı kategorilerden seçilmiş kelimeler gösterilmektedir.

Veri setindeki kelimelere ait görüntülerin özniteliklerinin çıkarımı aşamasında Ayrık Kosinüs Dönüşümü'nden yararlanılmıştır. Dönüşüm ile elde edilen, video görüntü ve derinlik haritalarına ait katsayıların belirli uzunlukta zig-zag tarama ile okunup 1:1 (bire bir) oranında birleştirilmesiyle öznitelik vektörleri oluşturulmuştur. Öznitelik vektörlerinde kullanılan DCT katsayılarının miktarının tanıma performansına etkisini gözlemleyebilmek adına, her işaret için 160×120, 320×240 ve 640×480 çerçeve boyutlarında uzunlukları sırasıyla 10-500 (10 artımlı), 40-2000 (40 artımlı) ve 160-8000 (160 artımlı) arasında değişen 50 adet öznitelik vektörü oluşturulmuştur. Farklı uzunluktaki öznitelik vektörlerinin tanımadaki

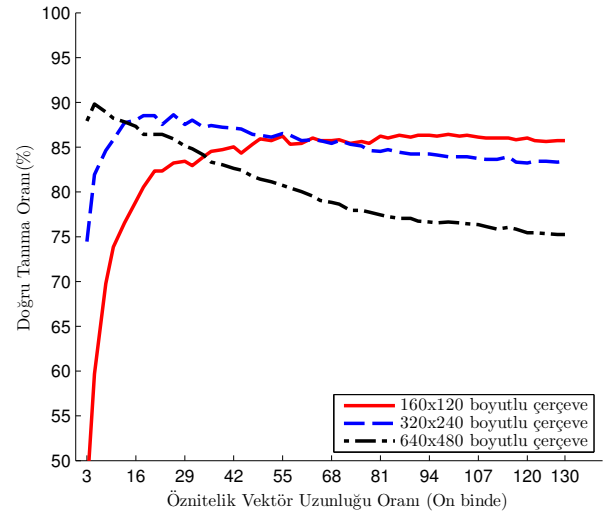
Tablo I. ÇALIŞMA KAPSAMINDA KULLANILAN 3 AYRI KATEGORİYE AIT TÜRK İŞARET DİLİ ÖRNEK KELİMELERİ

Kategori-1 Günlük konuşma	Kategori-2 Zaman terimleri	Kategori-3 Meslekler
günaydın	bazen	bankacı
iyi akşamlar	bugün	berber
iyi geceler	cuma	dişçi
merhaba	nisan	elektrikçi
özür dilerim	hafta	gazeteci
teşekkürler	yıl	mühendis
lütfen	zaman	öğrenci
sağol	sabah	öğretmen

performansları da bu sayede gözlemlenmiştir. Şekil 3'te farklı çerçeve boyutlarında, ilgili özvektörlerin doğru tanıma oranları verilmiştir.

Sistemdeki kelimeleri tanımak amacıyla Manhattan mesafesini kullanan en yakın komşu sınıflandırıcısı (K-NN) kullanılmıştır. Sistemin performansının değerlendirilmesinde veri seti örnekleri 3-katlı çapraz doğrulama ile işleme tabi tutulmuştur. Sistem performansının değerlendirilmesi aşamasında ilgili veri setinde doğru tanıma oranları (correct recognition rate - CRR) hesaplanmıştır. CRR hesaplanmasında doğru tanınan örneklerin sayısının veri setindeki toplam örnek sayısına oranı (3) alınmıştır.

$$CRR = \frac{\# \text{doğru tanınan işaret örnekleri}}{\# \text{tüm işaret örnekleri}} \quad (3)$$



Şekil 3. 1-en yakın komşuyu kullanan K-NN sınıflandırıcısının farklı çerçeve boyutlarında ve uzunluktaki öznitelik vektörlerinde başarımları (3-katlı çapraz doğrulama).

Şekil 3'te yer alan tanıma başarımlarına ait grafik analiz edildiğinde, 640×480 çerçeve boyutlarındaki görüntülerde tanıma oranı diğer boyutlara nazaran çok küçük katsayı oranlarında yüksek çıkmaktadır ancak katsayı miktarı arttıkça başarımlar belirgin bir şekilde düşmektedir. Düşük boyutlu çerçevelerde ise katsayı miktarı arttıkça başarımların arttığı görülmektedir. Bu sebeple işlem yükünü ve süresini azaltmak için

belirli bir başarımları oranı kaybı göze alınarak düşük boyutlu görüntülerde çalışmak tercih edilebilir.

Tablo II’de belirli oranlardaki DCT katsayılarında gri- imge ve derinlik haritalarının tüm veri setindeki başarımları gösterilmiştir.

Tablo II. MANHATTAN MESAFESİNİ KULLANAN KNN SINIFLANDIRICISININ FARKLI ÖZNETELİK VEKTÖR BOYUTLARINDA DOĞRU TANIMA ORANLARI (320×240 ÇERÇEVE BOYUTU)

Öznetelik Vektörü Uzunluğu (DCT)	DCT Katsayı Oranı (On binde ~)	Doğru Tanıma Oranları (%)	
		Gri-İmge	Derinlik
40	5	70,86	73,45
80	10	78,84	81,34
120	16	80,44	84,23
160	21	80,54	84,33
200	26	81,34	84,43
240	31	80,84	84,03
280	36	81,64	84,33
320	42	80,94	84,43
360	47	80,94	84,33
400	52	81,44	84,03

Veri setindeki kelimelerin tanınmasında, gri- imge ve derinlik haritalarına ait özneteliklerin ayrı ayrı ve birlikte değerlendirilmesiyle elde edilen doğru tanıma oranları tablo III’te sunulmuştur. Tablo III’deki tanıma oranlarına göre işaret diline ait hareketlerde derinlik bilgilerinin gri- imgeye göre daha başarılı olduğu görülmektedir. Gri- imge ve derinlik bilgilerinin birlikte kullanılması sistem başarımlarını daha da arttırmaktadır.

Tablo III. GRI-İMGE, DERİNLİK VE GRI-İMGE+DERİNLİK ÖZNETELİK VEKTÖRLERİNİN FARKLI ÇERÇEVE BOYUTLARINDA TANIMA PERFORMANSLARI (TÜM KATEGORİLER İÇİN - 111 KELİME)

	Doğru Tanıma Oranları (%)					
	160×120		320×240		640×480	
	DCT ^a	CRR	DCT	CRR	DCT	CRR
Gri- imge	73	79,44	36	81,64	13	83,33
Derinlik	104	81,44	34	84,53	8	85,53
Gri- imge+Derinlik	99	86,43	26	88,62	5	89,82

^aÖznetelik vektörlerindeki DCT katsayı yaklaşık oranı (On binde).

Çalışmada, kategorilerin ayrı ayrı değerlendirilmesiyle elde edilen başarımlar da gözlemlenerek tablo IV’te belirli bir DCT katsayı oranı ve çerçeve boyutu için gösterilmiştir.

Tablo IV. GRI-İMGE, DERİNLİK VE GRI-İMGE+DERİNLİK ÖZNETELİK VEKTÖRLERİNİN 320×240 ÇERÇEVE BOYUTLARINDA FARKLI KATEGORİLERDEKİ TANIMA PERFORMANSLARI (ON BİNDE ~10 DCT KATSAYI ORANI)

	Doğru Tanıma Oranları (%)		
	Kategori - 1 (22 kelime)	Kategori - 2 (58 kelime)	Kategori -3 (31 kelime)
Gri- imge	90,40	71,10	87,05
Derinlik	92,93	77,57	87,05
Gri- imge+Derinlik	94,44	82,13	89,93

V. SONUÇ

Yapılan çalışmada Türk İşaret Dili’ne ait örnekler, ardışıl görüntü çerçevelerinin farklarına dayalı toplamsal imge yaklaşımıyla başarılı olarak tanınabilmektedir. Kullanılan yöntemle, hareketlerin zaman uzayındaki değişimleri, DCT dönüşümü

ile uzamsal uzaya aktarılmış ve karmaşık olmayan bu yöntemle öznetelik vektörleri elde edilmiştir. Tanıma aşamasında yine basit bir yöntem olan KNN sınıflandırıcısı kullanılmıştır. Geliştirilen sistemde, görüntüler üzerinde nesne bölütlemesi (el, yüz vs.) işlemleri yapılmamaktadır. Bu sistem avantajlı olmakla birlikte hareketli arkaplana sahip görüntülerde verim düşüklüğü oluşturabilir ve böyle durumlarda nesne bölütleme yaklaşımına başvurulabilir. İşaretlere ait ardışıl görüntü çerçevelerindeki hareket değişimleri, küçük değişimler süzgeçlenerek kullanılmaktadır. Çalışmada kullanılan veri setinde tamamı işitme ve konuşma engelli işaretçiler tarafından gerçekleştirilen 111 kelimeye karşılık gelen video görüntüleri kullanılmıştır. Kelimelerin 3 farklı işaretçi tarafından 3’er kez tekrarı ile oluşturulan veri setinde bazı örnekler yüksek gürültü sebebi ile değerlendirme dışında tutulurken bazı işaretçilere ve kelimelere ait mevcut 4. örnekler performans ölçümü işlemlerine dahil edilmiş ve çalışmada toplam 1002 video kaydı kullanılmıştır. İşaretlere ait görüntüler Kinect duyarga üzerindeki RGB kamera ve kızılötesi algılayıcı kullanılarak 640×480 çerçeve boyutlarında kayda alınmıştır. İşaretlere ait öznetelik vektörleri, RGB kameradan alınan ve gri- imgelere dönüştürülen görüntülerle derinlik haritalarından elde edilmiştir. Aynı zamanda kullanılan yöntemle gri- imge ve derinlik haritalarında ayrı ayrı performans ölçümleri yapılmıştır. Elde edilen sonuçlarda işaretlere ait derinlik bilgilerinin tanıma etkin olduğu görülmüştür. Görüntü çerçeve boyutlarının tanıma performansına etkisini gözlemlemek amacıyla değişik çerçeve boyutları ile yapılan testlerde işlem karmaşıklığı ve tanıma başarımlarını göz önüne alarak küçük çerçeve boyutlarıyla (160×120) çalışmanın yeterli olduğu sonucuna varılmıştır. Sistemin performans değerlendirme aşamasında ise genel olarak %80 ile %90 arasında değişen başarı oranları gözlemlenmiştir.

TEŞEKKÜR

Bu çalışma, Yıldız Teknik Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından 2012-04-01-YL03 numaralı proje kapsamında desteklenmiştir.

KAYNAKÇA

- [1] C. Lucas, Ed., *The Sociolinguistics of Sign Languages*. Cambridge: Cambridge University Press., 2001.
- [2] C. Oz and M. C. Leu, “American sign language word with a sensory glove using artificial neural networks,” *Engineering Applications of Artificial Intelligence*, vol. 24, no. 7, pp. 1204–1213, Oct. 2011.
- [3] W. Gao, G. Fang, D. Zhao, and Y. Chen, “A chinese sign language recognition system based on SOFM/SRN/HMM,” *Pattern Recognition*, vol. 37, no. 12, pp. 2389–2402, Dec. 2004.
- [4] Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, “American sign language recognition with the kinect,” in *Proceedings of the 13th International Conference on Multimodal Interfaces*, Alicante, Nov. 14–18, 2011, pp. 279–286.
- [5] H. Haberdar and S. Albayrak, “Real time isolated turkish sign language recognition from video using hidden markov models with global features,” in *Computer and Information Sciences - ISCIS*, ser. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg, 2005, pp. 677–688.
- [6] T. Shanableh, K. Assaleh, and M. Al-Rousan, “Spatio-temporal feature extraction techniques for isolated gesture recognition in arabic sign language,” *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics*, vol. 37, no. 3, pp. 641–650, Jun. 2007.
- [7] O. Altun and S. Albayrak, “Turkish fingerspelling recognition system using generalized hough transform, interest regions, and local descriptors,” *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1626–1632, Oct. 2011.