

Konuşmacı Dönüştürme İçin Üç Yeni Yöntem

Oytun Türk

Levent M. Arslan

Elektrik-Elektronik Mühendisliği Bölümü, Boğaziçi Üniversitesi, İstanbul

Sestek A.Ş., AR-GE Bölümü, İstanbul

oytun@sestek.com.tr

arslanle@boun.edu.tr

Özetçe

Bu çalışmada konuşmacı dönüştürme için üç yeni yöntem geliştirilmiştir. İlk iki yöntem ses borusu özneliklerinin bir konuşmacıdan diğerine dönüştürülmesini amaçlamaktadır. Bu yöntemlerden ilki dalgacık dönüşümü, ikincisi ise seçimli önvurgulama kullanılarak ses işaretlerinin farklı sıklık aralıklarında işlenmelerine olanak sağlamaktadır. Üçüncü yöntem, ses perdesinin zamanla değişiminin konuşmacıya bağlı olarak modellenmesini ve dönüştürülmesini sağlamaktadır. Tüm yöntemlerin değerlendirilmesi için bir konuşmacı dönüştürme veri tabanı hazırlanmış ve öznel dinleme deneyleri tasarlanmıştır. Bu deneylerde, çeşitli konuşmacı dönüştürme yöntemleri, akustik öznelik nakilleri ve konuşmacıların özgün ses kayıtları karşılaştırılmıştır. Karşılaştırmalar hedef konuşmacıya benzerlik ve ses kalitesi açısından yapılmıştır. Alt-bant işleme dayalı ses borusu dönüştürme yöntemlerinin daha düşük doğrusal öngörü derecelerinde alt-bant işleme kullanılmayan yöntemle yakın kalitede çıktı sağladığı gösterilmiştir. Geliştirilen konuşmacıya bağlı entonasyon modelinin hedef konuşmacıya benzerliği arttırdığı gözlemlenmiştir.

1. Giriş

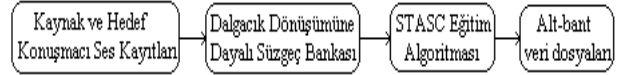
Ses borusu yapısının ve ses perdesinin zamanla değişiminin uygun şekilde modellenmesi ve dönüştürülmesi, konuşmacı dönüştürme algoritmalarının performansını doğrudan etkileyen bir faktördür. Ses borusu yapısı ve ses perdesi konuşmacı kimliğinin algılanmasında önemli iki etkidir [1]. Ses borusu spektrumunun modellenmesi ve dönüştürülmesi için formant frekansları [2], sinüzoidal model parametreleri [3] ve doğrusal spektrum frekansları [4] kullanılabilir. Doğrusal spektrum frekansları interpolasyon için uygun bir zemin oluşturduğundan konuşmacı dönüştürme uygulamalarında sıklıkla tercih edilmektedir [5]. Ses perdesinin zamanla değişimi, vurgunun önemli bir parçası olan entonasyonu belirlemektedir. Entonasyonun modellenmesi ve dönüştürülmesi için çeşitli yöntemler [6]'da karşılaştırılmıştır.

Bu çalışmada konuşmacı dönüştürme algoritmalarında kullanılabilecek üç yeni yöntem tasarlanmıştır. 2. bölümde ses borusu yapısı dönüşümü için dalgacık dönüşümüne dayalı bir yöntem geliştirilmiştir. 3. bölümde bant-geçiren süzgeçlemeye dayalı bir başka ses borusu spektrumu modelleme ve dönüşüm metodu açıklanmaktadır. 4. bölümde geliştirilen entonasyon modeliyle ses perdesinin zamanla değişimi ayrıntılı biçimde modellenip konuşmacılar arasında dönüştürülebilmektedir. Tüm yeni yöntemler STASC algoritması [4] çatısı altında birleştirilmiştir. 5. bölümde tasarlanan deneylerle yeni yöntemler ve STASC algoritması karşılaştırılmıştır. 6. bölümde sonuçlar tartışılmıştır.

2. Dalgacık Dönüşümüne Dayalı Alt-bant İşleme Sistemi

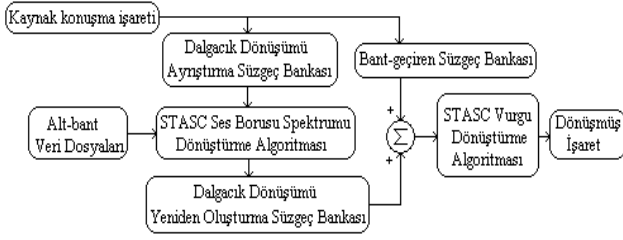
Bu bölümde, STASC konuşmacı dönüştürme algoritması [4], dalgacık dönüşümüne dayalı alt-bant işleme kullanılacak şekilde geliştirilmiştir. STASC algoritmasında kaynak konuşmacının sesinin hedef konuşmacının sesine dönüşümü iki adımda gerçekleştirilir: Eğitim ve dönüştürme. Eğitim aşamasında kaynak ve hedef konuşmacıdan içerikleri aynı olan ses kayıtları alınır. Her kaynak konuşmacı kaydı için bir Saklı Markov Modeli eğitilir. Bu model, hedef konuşmacı kaydının Viterbi algoritmasıyla etiketlenmesinde kullanılır. Kaynak ve hedef konuşmacının birbirine karşılık gelen akustik birimlerinden (fonemler ya da daha kısa ses parçaları) akustik parametreler (doğrusal spektrum frekansları, f0, süre, ve enerji) hesaplanır ve dönüşüm aşamasında kullanılmak üzere saklanır. Dönüşüm aşamasında ses borusu yapısı ve ses perdesi ayrı ayrı işlenerek çıktı işaretinde hedef konuşmacının ses özellikleri oluşturulur.

Geliştirilen dalgacık dönüşümüne dayalı eğitim yönteminin akış şeması Şekil 1'de gösterilmiştir. Bu yöntemde kaynak ve hedef konuşmacı kayıtları dalgacık dönüşümüyle alt-bant bileşenlerine ayrılır. Dönüştürülecek her alt-bant için STASC eğitim algoritmasıyla eğitim gerçekleştirilir. Kaynak ve hedef konuşmacının farklı alt-bantlarda birbirine karşılık gelen akustik parametreleri dönüştürmede kullanılmak üzere alt-bant veri dosyalarında saklanır.



Şekil 1: Dalgacık dönüşümüne dayalı eğitim.

Dönüştürme için akış şeması Şekil 2'de verilmiştir. Bu yöntemle istenen frekans aralıklarında farklı çözünürlükte dönüşüm gerçekleştirmek mümkündür. Yüksek örnekleme frekanslarında (44.1 KHz ve üzeri), yüksek frekans bölgeleri dönüştürülerek konuşma içermeyen spektrum bölgelerinden kaynaklanacak gürültü engellenebilir. Dalgacık dönüşümüyle elde edilen alt-bant işaretlerinin örnekleme frekansı, seyreltme işlemiyle azaldığından ses borusu spektrumunu daha düşük derecede Doğrusal Öngörü Analiziyle (DÖA) modellemek ve dönüştürmek mümkündür. Örneğin 44100 Hz'de kaydedilmiş bir işaret, iki dereceli bir dalgacık dönüşümü süzgeç bankasından geçirilerek her biri 11025 Hz örnekleme frekansında dört alt-bant bileşeni elde edilir. 44100 Hz'de ses borusu spektrumunun yeterli ayrıntıda modellenmesi için 48-50. dereceden DÖA gerekiyken, 11025 Hz örnekleme frekansında 14-16. dereceden DÖA'yle spektrum ayrıntılı biçimde modellenmektedir. Parametre sayısının azalması konuşmacı dönüştürme algoritmasının hızını da arttırmaktadır.



Şekil 2: Dalgacık dönüşümüne dayalı dönüştürme.

3. Seçimli Önvurgulamaya Dayalı Sistem

Konuşma işaretlerine DÖA öncesi önvurgulama uygulanması analizin nümerik özelliklerinin iyileştirilmesi açısından faydalıdır. Önvurgulamada amaç, konuşma işaretlerinde genelde düşük frekanslara göre daha düşük enerji seviyeli bileşenler içeren yüksek frekans bölgelerinin enerjilerinin artırılmasıdır. Bu şekilde yüksek frekans bölgeleri daha doğru biçimde modellenmektedir. Bu bölümde geliştirilen yöntem önvurgulamaya aynı temele dayanmaktadır. DÖA öncesi bant-geçiren filtrelerle elde edilen farklı frekans bantlarından tüm sinyalin DÖA spektrumu oluşturulmaktadır. Bu yöntem, seçimli önvurgulama olarak adlandırılmıştır. 2. bölümde açıklanan dalgacık dönüşümüne dayalı yöntem yüksek örnekleme frekanslarında az sayıda parametreyle ses borusu spektrumunun dönüştürülmesine olanak sağlamaktadır. Bu yöntemin en önemli dezavantajı, tüm alt-bant sinyalleri dönüştürüldüğünde örtüşme nedeniyle oluşabilen gürültüdür. Bunun engellenmesi için seyreltme uygulanmadan farklı frekans bölgelerini farklı çözünürlüklerde modelleyecek bir yöntem tasarlanmıştır. DÖA'yle spektrum tepeleri, dip noktalara göre daha iyi modellenmektedir. Bant-geçiren süzgeçler yardımıyla spektrumun belli bir aralığının vurgulanması ve DÖA'yle daha ayrıntılı modellenmesi mümkündür. Seçimli önvurgulama, Bölüm 3.1'de açıklanan analiz ve sentez aşamalarından oluşmaktadır.

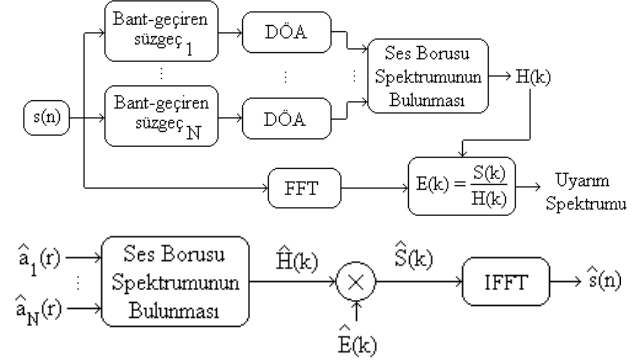
3.1. Analiz ve Sentez

Seçimli önvurgulamada ses borusu spektrumu, alt-bant bileşenlerinin spektral zarflarının ağırlıklı ortalamasıyla hesaplanmaktadır. Analiz sisteminin akış şeması Şekil 3'te gösterilmiştir. Konuşma işareti, $s(n)$, bant-geçiren bir süzgeç bankasından geçirilir. Her alt-bant bileşeni zamanda kaydırılan pencereler kullanılarak DÖA'yle incelenir ve i no'lu alt-bant için doğrusal öngörü katsayıları, $a_i(r)$, bulunur. Bu katsayılardan DÖA spektrumu, $H_i(k)$, hesaplanır. 1. denklem kullanılarak ses borusu spektrumu, $H(k)$, alt-bant spektrumlarından hesaplanır. Her alt-bant bileşeninin, spektrumun k indeksli örneğine katkısı 2. denklemde $c_i(k)$ ile gösterilmiştir. k_1 , $i+1$ no'lu alt-bantın düşük kesim frekansı, k_2 ise i no'lu alt-bantın yüksek kesim frekansıdır. $k_1 \leq k \leq k_2$ koşulu alt-bantların belli bir oranda çakışmasını sağlamaktadır.

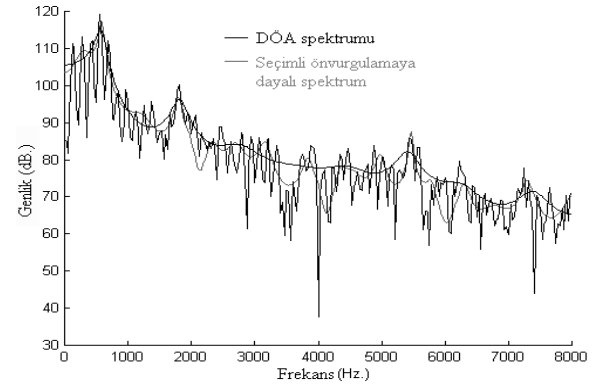
$$H(k) = \sum_i c_i(k) H_i(k) \quad (1)$$

$$c_i(k) = \begin{cases} \frac{k - k_1}{k_1 - k_2} + 1, & k_1 \leq k \leq k_2 \\ 0, & k < k_1 \text{ veya } k_2 < k \end{cases} \quad (2)$$

Sentez aşamasında, sentez DÖA katsayıları ve sentez uyarım spektrumu kullanılarak çıktı işareti spektrumu elde edilir. Sentez parametreleri uygulamaya bağlı olarak analiz parametrelerinin değiştirilmiş hali olabilir. Şekil 3'te sentez yönteminin akış şeması gösterilmiştir. Şekil 4'te seçimli önvurgulama yöntemine bir örnek verilmiştir. Analiz ve sentez için 0.0-8.0 KHz aralığını 4 eşit alt-banda bölen bir bant-geçiren süzgeç bankası kullanılmıştır. Tüm-bant DÖA'nde ve seçimli önvurgulama yönteminde aynı öngörü derecesi (18) kullanılmıştır. İşlenen konuşma işaretinin örnekleme frekansı 16 KHz'dir.



Şekil 3: Seçimli önvurgulama analiz (üst) ve sentez (alt) algoritmaları akış şemaları.



Şekil 4: DÖA ve seçimli önvurgulamaya dayalı spektral kestirim.

Frek. Aralığı (Hz.)	Seçimli Önvurgulama		DÖA	
	Ortalama	Std. Sap.	Ortalama	Std. Sap.
0-22050	0.43	0.10	0.44	0.11
0-1034	1.00	1.11	1.05	1.18
1034-1895	0.45	0.30	0.54	0.40
1895-2756	0.44	0.24	0.49	0.34
2756-3618	0.40	0.22	0.44	0.26
3618-4823	0.39	0.18	0.44	0.21
4823-6546	0.37	0.12	0.40	0.13
6546-8269	0.38	0.12	0.40	0.13
8269-11714	0.39	0.08	0.39	0.08
11714-15159	0.39	0.08	0.40	0.09
15159-22050	0.42	0.07	0.40	0.07

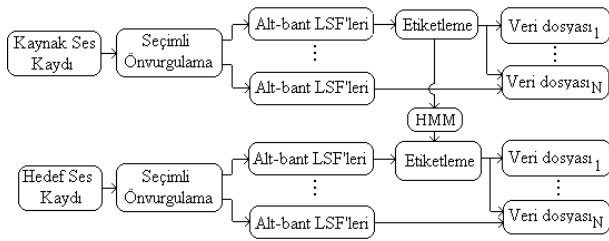
Tablo 1: 50. derece DÖA ve 24. derece seçimli önvurgulamanın spektral mesafeler yönünden karşılaştırılması

DÖA ve seçimli önvurgulama yöntemlerinin spektral kestirim performansının karşılaştırılması için yapılan nesnel deneyin sonuçları Tablo 1'de verilmiştir. Bu deneyde, 44.1 KHz örnekleme frekansında kaydedilmiş konuşma işaretlerinde kestirilen spektrum'un Fourier spektrumuna uzaklığı hesaplanmıştır. Tablo 1'de seçimli önvurgulamanın

tüm-bant DÖA'ya göre daha düşük öngörü derecesinde daha iyi sonuç verdiği görülmektedir.

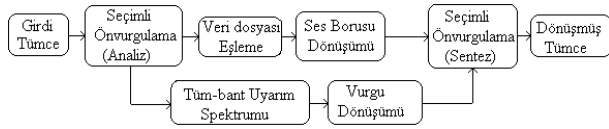
3.2. Eğitim ve Dönüştürme

Seçimli önvurgulama yöntemi STASC konuşmacı dönüştürme algoritmasına [4] entegre edilmiştir. Seçimli önvurgulama için 10 sonlu dürtü yanıtı süzgeçten oluşan bir süzgeç bankası tasarlanmıştır. Eğitim aşamasında kaynak ve hedef konuşmacının aynı içerikli ses kayıtları seçimli önvurgulamayla analiz edilir. Kaynak konuşmacı işaretinin en düşük frekans aralığını kapsayan alt-bant bileşeni için bir Saklı Markov Modeli eğitilir. Hedef konuşmacı işareti aynı modelle etiketlenir. Her alt-bantta birbirine karşılık gelen kısımlar için ses borusu parametreleri (doğrusal spektrum frekansları) hesaplanır. Bu parametreler alt-bant veri dosyalarında saklanarak dönüşüm sırasında kullanılır. Şekil 5'te seçimli önvurgulamaya dayalı eğitim yönteminin akış şeması gösterilmiştir.



Şekil 5: Seçimli önvurgulamaya dayalı eğitim.

Her alt-bant ses borusu spektrumunun dönüştürülmesi için alt-bant veri dosyaları kullanılır. Bu amaçla, kaynak işareti seçimli önvurgulamayla analiz edilir. Tüm-bant uyarım spektrumu vurgu dönüşümü için ayrıca işlenir. Her alt-bant ses borusu bileşeni veri dosyalarındaki en yakın hedef parametrelerinin ağırlıklı ortalaması alınarak dönüştürülür. Dönüştürülmüş alt-bant bileşenleri Bölüm 3.1'de açıklanan seçimli önvurgulama sentez yöntemiyle birleştirilir. Seçimli önvurgulamaya dayalı dönüştürme algoritmasının akış şeması Şekil 6'da verilmiştir.



Şekil 6: Seçimli önvurgulamaya dayalı dönüştürme.

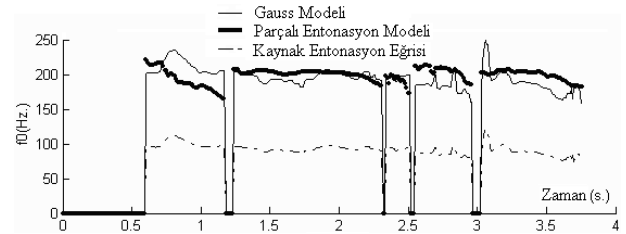
4. Parçalı Entonasyon Modeli

Entonasyonun modellenmesi için sıklıkla kullanılan yöntem, f0 değerlerinin olasılık dağılım işlevinin bir Gauss dağılımı olduğunun varsayılmasıdır [4], [6]. Bu bölümde entonasyonun daha ayrıntılı olarak modellenmesi ve dönüştürülmesi için yeni bir yöntem olan parçalı entonasyon modeli açıklanmaktadır. Öncelikle kaynak ve hedef konuşmacının aynı içerikli ses kayıtları etiketlenir. Entonasyon eğrileri özilinti işlevine dayalı bir yöntemle hesaplanır. Eğrilerin sesli parçaları belirlenir. Hedef konuşmacı ses perdesi eğrilerindeki sessiz kısımlar komşu sesli kısımların değerleri kullanılarak interpolate edilir. Bunun nedeni kaynak konuşmacıda sesli olan kısımların hedef

konuşmacıda sessiz kısımlara denk gelebilmesi ve bu durumun dönüşümde gürültüye neden olmasıdır. Etiketleme bilgisi kullanılarak birbirine karşılık gelen kaynak ve hedef konuşmacı entonasyon eğrisi parçaları bulunur. 3. denklemde $s_i(n)$, i no'lu kaynak entonasyon eğrisi parçasının n no'lu f0 değerini, $t_i(n)$ ise karşılık gelen hedef ses perdesi parçasının n no'lu f0 değerini göstermektedir. Dönüşüm aşamasında girdi entonasyon eğrisinin sesli kısımları, $f_j(n)$, bulunur. j no'lu sesli kısım için girdi, kaynak ve hedef ses perdesi eğri parçaları aynı uzunluğa (N_j) interpolate edilir. Kaynak entonasyon eğri parçalarının girdi entonasyon eğrisi parçasına uzaklığı, d_i , 3. denklemle hesaplanır. 4. denklemde her hedef entonasyon eğri parçası için d_i 'ye bağlı bir ağırlıklı hesaplanır. Bu ağırlıklar, her hedef entonasyon eğri parçasının ağırlıklı ortalamasının alınarak çıktığı f0 değerleri, $o_j(n)$ 'in, bulunmasında kullanılır (5. denklem). Şekil 7'de parçalı entonasyon modeliyle bir entonasyon eğrisi dönüşüm örneği verilmiştir.

$$d_i = \frac{\sum_{n=1}^{N_j} |f_j(n) - s_i(n)|^2}{\sum_{all i} \sum_{n=1}^{N_j} |f_j(n) - s_i(n)|^2} \quad (3)$$

$$w_i = \frac{\exp(-\alpha d_i)}{\sum_{all i} \exp(-\alpha d_i)} \quad (4) \quad o_j(n) = \sum_{all i} w_i t_i(n) \quad (5)$$



Şekil 7: Parçalı entonasyon modeliyle f0 dönüşümü.

5. Değerlendirme

Bu bölümde, üç farklı ses borusu spektrumu dönüştürme ve iki farklı entonasyon dönüştürme yönteminin karşılaştırılması için özel bir dinleme deneyi tasarlanmıştır. Ses borusu dönüştürme yöntemleri STASC [4], dalgacık dönüşümüne dayalı yöntem (DD) ve seçimli önvurgulama (SÖ) yöntemleridir. Entonasyon dönüşümü için Gauss modeliyle (GM) parçalı entonasyon modeli (PEM) kullanılmıştır. 4 erkek, 4 kadın konuşmacıdan 30 tümce ve 50 sözcük içeren Türkçe bir veri tabanı kaydedilmiştir. Kayıtların örnekleme frekansı 44.1 KHz'dir. Öncelikle STASC, DD, SÖ, GM ve PEM yöntemlerinin eğitim aşamaları her kaynak/hedef konuşmacı ikilisi için gerçekleştirilmiştir. 5 test tümcesi ve 10 test sözcüğü Tablo 2'de gösterilen tüm olası kombinasyonlar kullanılarak dönüştürülmüştür. Karşılaştırma için ses borusu yapısı ve ses perdesi nakilleri de kullanılmıştır. Nakiller herhangi bir akustik özneliliğin dönüşümü için ideal duruma karşılık gelmektedir.

Toplam 10 deneye 112 ses kayıt üçlüsü dinletilmiştir. Üçlülerdeki birinci ve ikinci kayıt kaynak ve hedef konuşmacıların özgün ses kayıtlarıdır. Üçüncü kayıt denek tarafından değerlendirilecek çıktılarıdır. Biridir: özgün bir kayıt, akustik öznelilik nakli (Tablo 2'nin 1. ve 2. sırası), dönüşüm çıktısı (Tablo 2'nin 3-11. sıraları). Deneklerden üç skor vermeleri istenmiştir: Ses Kimliği, kimlik için verilen karara olan güven ve Ses Kalitesi. Kimlik skoru "Kaynak",

“Arada” ve “Hedef” seçimlerinin sırasıyla 0.0, 0.5 ve 1.0’e eşlenmesiyle oluşturulmuştur. Güven skoru, deneğin kimlik skorunu ne kadar güvenle verdiğini ölçmektedir. Kalite skoru, çıktıların özgün kayıtlara göre kalitesini göstermektedir. Güven ve kalite skoru 1 (yüksek) ile 5 (düşük) arasındaki tamsayılardan oluşmaktadır. Deneklerden, çıktının üçüncü bir şahısa benzemesi durumunda kimlik skoru olarak 0.5, güven skoru olarak en düşük skor olan 1’i seçmeleri istenmiştir. Skorlar en yüksek değerleri 1.0 olacak şekilde normalize edilmiştir. Tüm skorlar için ortalama değer ve çeyreklerarası aralığı (CAA) değerleri hesaplanmıştır. Sonuçlar Şekil 8’de gösterilmektedir. Şekil 7’deki grafiklerin her birindeki gruplar farklı cinsiyet kombinasyonlarında elde edilen sonuçları göstermektedir. Örneğin E→K’da kaynak konuşmacı erkek, hedef konuşmacı kadındır. “Tümü” durumunda skorlar tüm cinsiyet kombinasyonları için hesaplanmıştır. Her grupta Tablo 2’deki kombinasyonlara karşılık gelen 11 çizgi gösterilmiştir.

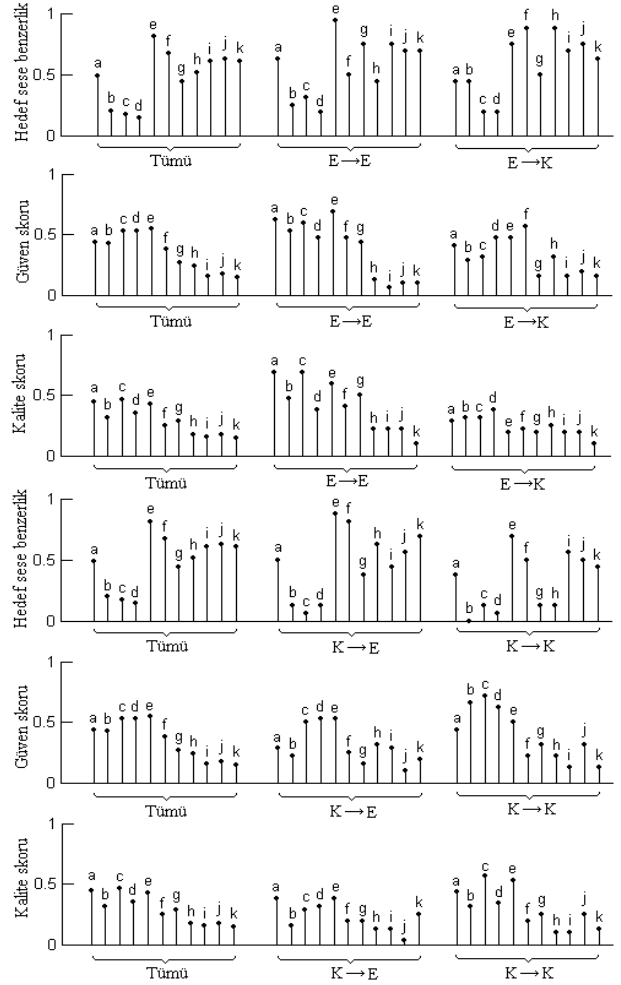
SB Dönüşümü	P Dönüşümü	Sembol	Çıktı Türü
-	-	a	SB Nakli
STASC	-	b	SB Dönüşümü
DD	-	c	SB Dönüşümü
SÖ	-	d	SB Dönüşümü
-	-	e	SB+P Nakli
STASC	GM	f	SB+P Dönüşümü
STASC	PEM	g	SB+P Dönüşümü
DD	GM	h	SB+P Dönüşümü
DD	PEM	i	SB+P Dönüşümü
SÖ	GM	j	SB+P Dönüşümü
SÖ	PEM	k	SB+P Dönüşümü

Tablo 2: Test edilen konuşmacı dönüştürme yöntemleri.
(SB: Ses borusu, P: Ses perdesi)

Şekil 8’de sadece ses borusu yapısını dönüştürmenin hedef konuşmacıya yeterli benzerliği sağlamadığı görülmektedir. Ses borusu naklinde bile çıktı kaynak konuşmacıyla hedef konuşmacı arasında olarak değerlendirilmiştir. Ses perdesi de dönüştürüldüğünde hedef konuşmacıya benzerlik artmaktadır. İşleme miktarı arttıkça güven ve kalite skorları düşmektedir. Farklı ses borusu dönüştürme yöntemleri için kaynak ve hedef konuşmacı cinsiyetlerine bağlı olarak farklı sonuçlar elde edilmiştir. STASC yöntemi farklı cinsiyet kombinasyonlarına diğer yöntemlerden daha uyumludur. Parçalı entonasyon modeli hedef konuşmacıya benzerliği arttırmıştır.

6. Tartışma

Bu çalışmada ses borusu sepkrumu dönüşümü için iki, entonasyon dönüşümü içinse bir yeni yöntem geliştirilmiştir. Ses borusu dönüşüm yöntemlerinde alt-bant işleme metotları kullanılarak DÖA’ne göre daha az parametreyle ayrıntılı ses borusu spektrumu kestirimi gerçekleştirilmiştir. Bu yöntemler genel olarak öngörü derecesinin artırılmasına karşılık gelmektedir. Yüksek örnekleme frekanslarında öngörü derecesinin fazla artırılması mümkün olmadığından, yeni yöntemler kullanılarak ayrıntılı modelleme ve dönüşüm yapılabilir. Özellikle film seslendirme uygulamalarında yüksek örnekleme frekansları (>44.1 KHz) tercih edilmektedir. Bu uygulamalarda kullanılacak konuşmacı dönüştürme metotları için alt-bant işlemeye dayalı yöntemler uygun bir zemin oluşturmaktadır.



Şekil 8: Öznel deney sonuçları

7. Kaynakça

- [1] Türk, O. *New Methods For Voice Conversion*. Yüksek Lisans Tezi. Boğaziçi Üniversitesi, 2003.
- [2] Gutierrez-Arriola, J.M., Hsiao, Y.S., Montero, J.M., Pardo, J.M., ve Childers, D.G., “Voice Conversion Based On Parameter Transformation”, *Proc. of the ICSLP 1998*, Vol. 3, sf. 987-990, Sidney, Avustralya.
- [3] Stylianou, Y., Cappe, O., ve Moulines, E., “Continuous Probabilistic Transform for Voice Conversion”, *IEEE Transactions on Speech and Audio Processing*, Vol. 6, No. 2, 1998, sf. 131-142.
- [4] Arslan, L.M., “Speaker Transformation Algorithm Using Segmental Codebooks”, *Speech Communication* 28 (1999), sf. 211-226.
- [5] Kain, A.B., ve Macon, M., “Personalizing A Speech Synthesizer by Voice Adaptation”, *Proc. of the 3rd ESCA/COCOSDA International Speech Synthesis Workshop*, 1998, sf. 225-230.
- [6] Chappell, D.T., ve Hansen, J.H.L., “Speaker-Specific Pitch Contour Modeling and Modification”, *Proc. of the ICASSP 1998*, Vol. II, sf. 885-888, Seattle, ABD.
- [7] Türk, O., ve Arslan, L.M., “Subband Based Voice Conversion”, *Proc. of the ICSLP 2002*, Vol. 1, sf. 289-292, Denver, Colorado, ABD.