

Türkçe Metinden Konuşma Sentezlemede Bugünkü Durum

II. Bölüm

Elektronik Mühendisi İbrahim Baran Uslu
ibuslu@baskent.edu.tr

EMO Ankara Şubesi Bülteninin 2010/2 nolu sayısında yayınlanan ilk yazımızda, Metinden Konuşma Sentezleme (MKS) sistemlerinden genel olarak bahsetmiştik. Bu yazımızda ise, Türkçe MKS sistemleri üzerine yapılmış olan çalışmaları ve bugünkü durumu inceleyeceğiz. EMO'nun 1993 yılında düzenlediği 5. Ulusal EEBM Kongresine katılan Mete BULUT (Doç. Dr., EEM, ODTÜ), o yılki makalesinde; geliştirdikleri Türkçe MKS sistemini anlatmıştır (1). Donanım olarak gerçekleştirilen bu sentezleyici, sesbirimleri (fonem) birleştirmektedir. Yazımızın amacı o günden bugüne, Türkçe MKS sistemleri konusunda alınan yolu ortaya koymaktır.

Giriş

Türkçe, Ural-Altay dilleri grubuna bağlı eklemeli bir dildir (2). İngilizcedeki gibi çok sayıda telaffuz (pronunciation) farklılıkları içermemesi bir avantajdır. Örnek; İngilizcedeki satranç anlamındaki (chess) sözcüğünde, 'ch' difonu (ikili sesbirim), 'ç' ile sesletilirken; koro anlamındaki (chorus) sözcüğündeki aynı difon, 'k' ile sesletilir. Ancak, Türkçemizde bazı sesletim farklılıkları da yok değildir. Örnek; âile (â: uzun a), kâğıt (â: ince a), lâle (â: uzun ve ince a), ihmâl (â: yumuşak a) sözcüklerindeki a sesi, 'sana' derken kullanılan a'dan farklıdır (3). Bu sözcüklerin aynı a ile sesletilmesi güzel bir konuşma sentezi vermeyecektir. Ayrıca Türkçede 'ʌ' düzeltme işaretinin kaldırıldığı da doğru olmayıp, son 20-25 yıldır var olan asılsız bir söylentidir (3). Yazı dilinden çıkartılan istatistikler, konuşma diline yönelik ipuçları verebilir. Türkçedeki harf, sesbirim, difon, trifon gibi ses parçalarının dağılımları incelenmiştir (4, 5). Buradan elde edilen sonuçlar, konuşma dilinin daha hızlı ve kolay kapsanmasını sağlayabilir.

Lisansüstü Çalışmalar

Türkçe metinden konuşma sentezlemede ve ilgili konularda, bugüne kadar yapılmış olan lisansüstü tez çalışmaları Çizelge-1'de verilmiştir. Bu çalışmalar incelendiğinde, genel olarak eklemeli sentezleyiciler üzerinde durulduğu, birleştirilen parçaların önceleri sesbirim, daha sonra hece ve son çalışmalarda çoğunlukla difonlar olduğu görülmektedir. Ayrıca, eklemeli sentezleyicilerde, farklı sinyal işleme (birleştirme) yöntemlerinin (değişik örtüştürüp ekleme yöntemleri, sinüzoidal model gibi) başarıları incelenmiştir. Son yapılan çalışmaların büyük bir bölümünde ise, daha doğal bir konuşma sentezi için, süre ve ezgi modelleri geliştirilmeye çalışılmıştır. Bir diğer hedefin, taşınabilir cihazlar için konuşma sentezleyici geliştirilmesi olduğu da not edilmelidir. Ses dönüştürme üzerine yapılan çalışmalar da bu çizelgede yer almıştır. Çünkü, geliştirilen sentezleyici ancak kayıt alınan konuşmacının sesiyle konuşabilir. Herkesin sesiyle kayıt almak zahmetli olacağı için, konuşmacıların seslerini birbirine dönüştürme amaçlı çalışmalar yapılmaktadır.

Yazar	Çalışmanın Türü, Yılı	Üniversite	Çalışmanın Başlığı
İlhan Yaşar ÖZÜM	Yüksek Lisans, 1993	ODTÜ	"A Speech Synthesis System for Turkish Language Based on the Concatenation of Phonemes Taken From a Speaker"
Kâmil GÜVEN	Yüksek Lisans, 1994	Çukurova	"PC Based Speech Synthesis for Turkish"
Murat Servet ERER	Yüksek Lisans, 1994	İTÜ	"Karma Söz Üretme Yöntemi ile Türkçe Yazılı Metinden Söze Geçme"
Kerem AYHAN	Yüksek Lisans, 1998	ODTÜ	"Text to Speech Synthesizer in Turkish Using Non Parametric Techniques"
Özgül SALOR	Yüksek Lisans, 1999	ODTÜ	"Signal Processing Aspects of Text to Speech Synthesizer in Turkish"
Barış BOZKURT	Yüksek Lisans, 2000	Boğaziçi	"Reading Aid for Visually Impaired (A Turkish Text-to-Speech System Development)"
Çağla ÖNÜR	Yüksek Lisans, 2001	ODTÜ	"Concatenative Speech Synthesis Based on a Sinusoidal Speech Model"
Erkan ABDULLAHBEŞE	Yüksek Lisans, 2001	Boğaziçi	"Fundamental Frequency Contour Synthesis for Turkish Text to Speech"
Şifa Serdar ÖZEN	Yüksek Lisans, 2002	Hacettepe	"Türkçe Metinden Konuşma Sentezleme"
Barış EKER	Yüksek Lisans, 2002	Bilkent	"Turkish Text to Speech System"
Ömer ŞAYLI	Yüksek Lisans, 2002	Boğaziçi	"Duration Analysis and Modeling for Turkish Text-to-Speech Synthesis"
Banu OSKAY	Yüksek Lisans, 2002	ODTÜ	"Automatic Modelling of Turkish Prosody"
Esra VURAL	Yüksek Lisans, 2003	Sabancı	"A Prosodic Turkish Text-to-Speech Synthesizer"
Oytun TÜRK	Yüksek Lisans, 2003	Boğaziçi	"New Methods for Voice Conversion"
Haşim SAK	Yüksek Lisans, 2004	Boğaziçi	"A Corpus Based Concatenative Speech Synthesis System for Turkish"
Ozan Aktan	Yüksek Lisans, 2004	Boğaziçi	"A Single Chip Solution for Text-to-Speech Synthesis"
Asude KARLI	Yüksek Lisans, 2005	Ankara	"Örnek Bir Dizi Cümle İçin Türkçe Metinden Konuşma Sentezleyici"
Özgül SALOR	Doktora, 2005	ODTÜ	"Voice Transformation and Development of Related Speech Analysis Tools for Turkish"
Özlem ÖZTÜRK	Doktora, 2005	ODTÜ	"Modeling Phoneme Durations and Fundamental Frequency Contours in Turkish Speech"
İlker ÜNALDI	Yüksek Lisans, 2007	Hacettepe	"Taşınabilir Cihazlar İçin Türkçe Metinden Konuşma Sentezleme Sistemi"
Oytun TÜRK	Doktora, 2007	Boğaziçi	"Cross Lingual Voice Conversion"

Çizelge-1: Türkçe MKS ve onunla ilgili konularda yapılan lisansüstü tez çalışmaları (tarihsel sıralı)

Sonuçlar ve Tartışma

Bugün itibariyle, elde edilen konuşmanın kalitesi endüstriyel bazı uygulamaların hayata geçirilmesine izin vermiştir. Bir gazetenin web sayfasına "sesli haber okuma" özelliği eklenmiş, ayrıca İDO (İstanbul Deniz Otobüsleri) işletmesi de sefer anonslarını ve bazı duyuruları konuşma sentezleyici kullanarak yapmaya başlamıştır (6). Ama hâlen yapılması gereken iyileştirmeler bulunmaktadır. Sentezleme kurallarının otomatik veya yarı-otomatik duruma getirilmesi amaçların başında gelmektedir. Ayrıca, birleştirilen ses parçaları için süre; sözcük ve tümceler için vurgu ve ezgi modelleri geliştirme araştırılan konular arasındadır.

Referanslar:

- (1) Bulut, M., "Türkçe metinden konuşma sentezleyici", EMO 5. Ulusal EEBM Kongresi, 210-215, 1993, Trabzon.
- (2) Ergin, M., Türk Dil Bilgisi, s. 9, Bayrak Basım, 1998.
- (3) Yılmaz, A. E., "Türkçe metinden konuşma sentezleme uygulamaları için bir veri sözlük seti ve yazılım çerçevesi", Gazi Üniv. Müh. Mim. Fak. Der., Cilt 24, No: 4, 735-744, 2009.
- (4) Salor, Ö., Pellom, B.L., Çiloğlu, T., Demirekler M., "Turkish speech corpora and recognition tools developed by porting SONIC: Towards multilingual speech recognition", Computer Speech and Language, vol. 21, pp. 580-593, 2007.
- (5) Uslu, İ.B., Yılmaz, A.E., İlk, H.G., "Statistical digram and trigram analysis of Turkish in terms of coverage and entropy for possible language and speech based applications", 23-27 Ağustos EUSIPCO 2010 için kabul edildi.
- (6) <http://www.sestek.com.tr/referanslar.html>