

第3世代携帯電話向けの分散型音声認識システムの実用化

受賞業績 **第3世代携帯電話における大語彙連続音声認識を可能にした分散型音声認識システムの開発と商用化**

加藤 恒夫^{*1} 宇都宮 栄二^{*1} 内藤 正樹^{*1} 内部 利明^{*2} 河井 恒^{*3}

^{*1}(株) KDDI 研究所 ^{*2}KDDI (株) ^{*3}情報通信研究機構

このたびは喜安記念業績賞を受賞することになり、大変光栄に存じます。改めて本技術の実用化に携わった関係者、音声認識技術を育てられた諸先輩方に感謝いたします。

受賞した第3世代携帯電話向けの分散型音声認識技術は、携帯電話とサーバ型音声認識が処理を分担する方式に基づくもので、2006年1月に「声de入力」機能としてリリースされました。以来、本機能は主要なコンシューマ向け携帯アプリである乗換検索、歩行者用ナビゲーション、同乗者用ナビゲーションなどに利用されています。本機能を搭載した携帯電話はこれまでに10メーカー150機種に及びます。

研究開発の発端は、2000年以降急速に普及した携帯インターネットや携帯アプリ向けに端末性能によって制限されない音声入力機能を提供できないかという単純な問いに対して、ちょうどその頃に標準化された分散型音声認識用の音響特徴量を利用し、データ通信に基づく分散システムを試作したことに始まります。分散型音声認識システムでは、高い認識精度に加えて、応答時間を端末・サーバ間の通信時間を含めて数秒以内に抑えることが求められます。携帯電話端末上の音響分析処理・ベクトル符号化をリアルタイム処理化する方式、サーバ型音声認識における効率的な探索方式などの考案により、応答時間を大幅に短縮し、実用的な精度と応答時間を両立させました。また、誤認識の可能性が高い場合には推定された誤認識原因をユーザに通知して誤認識の

繰り返しを防ぐ機能を加えるなど、音声認識のユーザビリティを高めました。

現在では、スマートフォンから利用できる超大語彙の音声認識が現れています。私達もこれまでに培った技術を基礎として、新たな音声アプリケーションの創出に寄与すべく努力していく所存です。

(2011年5月17日受付)

加藤 恒夫(正会員) tkato@kddilabs.jp

1996年東京大学大学院工学系研究科修士課程修了。KDDI研究所にて音声認識システムの研究開発に従事。2009年より同社ユーザインタフェースグループ グループリーダー。

宇都宮 栄二 ei-utsunomiya@kddilabs.jp

KDDI研究所にて、ソフトウェア開発支援技術や分散オブジェクト技術などの研究開発を経て、音声信号処理の技術開発・サービス開発に従事。現在、同社サービス開発グループ グループリーダー。

内藤 正樹 naito@kddilabs.jp

1991年電気通信大学大学院電気通信学研究所博士前期課程修了。1991～97年KDD研究所、1997～2000年ATR音声翻訳通信研究所、2000年よりKDDI研究所において音声認識システムの研究・開発に従事。博士(工学)。

内部 利明 to-uchibe@kddi.com

1994年東京工業大学大学院総合理工学研究科修士課程修了。KDDI研究所における音声認識、話者認識の研究開発を経て、現在KDDIにて電話サービスに関するソリューション業務に従事。

河井 恒 hisashi.kawai@nict.go.jp

1989年東京大学大学院工学系研究科博士課程修了。工学博士。1989～2000年KDD研究所、2000～04年ATR、2004～09年KDDI研究所、2009年より情報通信研究機構に勤務。音声の認識・合成、音響信号処理、音声翻訳の研究開発に従事。

