

JEITA 一般社団法人電子情報技術産業協会

平成24年度音声認識・合成技術応用製品の動向について

- 音声入出力方式標準化専門委員会実施の製品動向調査結果より -

2013 年 12 月

音声入出力方式標準化専門委員会

○概要

音声認識・合成技術の適用分野の拡大，応用製品の普及を目的として，本委員会では，継続して製品動向調査を実施している．本調査結果は，応用製品の開発者，製品の導入・購入を考える企業・個人に有益な情報を提供するとともに，継続実施することにより，技術の進歩，利用領域の変遷を理解する貴重な資料となっている．

平成 22 年度まではシステム開発者向け製品とエンドユーザー向けの製品の大きく 2 種類に分類し調査を行っていたが，主にコンシューマを対象にしたエンドユーザー向けの製品については，企業ばかりでなく個人・団体等からの提供も増えつつあり，市場・製品動向の網羅的な把握と調査が困難になってきたため平成 22 度をもって取りやめとし，本年度の調査は平成 22 年度以降の調査方法を基本的に踏襲し，以下の方法により実施した．

(1) 情報提供の依頼

平成 24 年 12 月に，JEITA 本委員会ホームページに，情報提供の依頼を掲示するとともに，音声研究・開発者向けメイリングリスト（onsei-mail@m.aist.go.jp）に同一内容のメールを配信．なお，前年に情報を提供いただいた企業には，直接メールを発信し，依頼した．

(2) 掲載製品

情報提供企業が製造あるいは販売する音声認識・合成関連製品であり，平成 24 年 12 月末現在で，日本国内で購入可能な製品とした．

(3) 提供内容

委員会で作成した，製品分類ごとの製品特徴・仕様を記載するフォームを配布，記載いただき，電子ファイルとして受領した．

(4) 掲載内容

提供いただいた情報は基本的に一切の加工を行わず掲載している．ただし，内容の変更を伴わない書式の統一，情報提供者の明らかな誤解による分類誤りについては，委員会にて変更をしている．また調査結果は，上記手法による本委員会からの依頼に対して回答のあった情報のみを掲載しているため，網羅性の視点からは完全な調査ではない．よって，周知であるにも関わらず調査結果に記載されていない製品が存在する．H23 年度は回答があったが H24 年度は回答が得られなかった機関については，製品自体が継続して販売されていることが企業ホームページから確認できる場合でも，該当製品情報を掲載していない．

記載されている内容については，委員会にて確認をしておらず，内容は，すべて情報提供企業にその責任がある．

○音声合成の製品動向調査結果

日本国内で販売されている音声合成製品を調査した結果について以下に概要をまとめる。

当製品調査は、調査対象の製品分類を大きく「ミドルウェア組み込み向け」、「ミドルウェア PC」、「ハードウェア」、「その他」に分けた上で（第 1 分類）、さらに細分類（第 2 分類）を設けて製品情報をまとめている。調査の概要で述べた通り平成 23 年度からコンシューマ向け製品を調査対象から除外したことで対応製品が存在しなくなったため、システム開発者向け製品のみを対象としている。調査結果を付録 1 表 A-1～9 として添付する。以下では、その調査結果をもとに、分類ごとの件数とその推移について述べる。

表 1 に、12 の細分類（第 2 分類）ごとの製品件数と推移を示す。製品の総数は計 115 件という結果であった。今回、3 項目については該当製品の報告がなかった。

表 1 分類毎の音声合成製品の件数と推移

分類		表番号*	H24		H23	H22
第 1 分類	第 2 分類		件数	増減**	件数	件数
ミドルウェア 組み込み向け	規則音声合成	表 A-1	36	+14	22	24
	録音・再生		0	0	0	0
	その他	表 A-2	1	0	1	0
ミドルウェア PC	規則音声合成	表 A-3	26	0	26	23
	録音・再生		0	0	0	0
	その他	表 A-4	2	+2	0	0
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	2	0	2	1
	録音・再生 (録音再生 LSI)		0	0	0	0
	その他	表 A-6	1	0	1	2
その他	サーバ製品	表 A-7	13	-5	18	8
	PC ソフトウェア・ツール等	表 A-8	25	+2	23	18
	サービス	表 A-9	10	+3	7	3
合計			116	16	100	79

*表番号：付録 1 製品一覧の対応表番号、 **増減：H21 年度からの増減

今年度の回答製品数は総計 116 件となり、平成 23 年度と比較して 16 件の増加となった。以下、平成 23 年度と比較した今年度の件数の推移をまとめる。

細分類で見ると、ミドルウェア組み込み向けの規則音声合成とミドルウェア PC では増加し、ハードウェア向け録音・再生、PC ソフトウェア・ツール等は同等であった一方、サーバ製品数は減少した。第 2 分類の 12 項目のうち、回答があった 10 項目について、調査結果の詳細については音声合成関連製品一覧の表 A-1～A-9 に掲載する。

○音声認識の製品動向調査

本年度調査にあたって、その製品分類・記載項目について検討し、昨年同様の分類ならびに記載項目で調査を実施することとした。具体的には、製品分類は音声認識製品を使用する立場から、①システム開発者向けとして、ミドルウェア（組込み向け）とミドルウェア（PC 向け）、②ソリューション利用者向けとして、CTI ソリューション、Web サービスソリューション、その他ソリューション、そして③エンドユーザ向けとして、応用製品とソフトウェア製品の、計 7 つに分類している。各分類における平成 22 年度～24 年度の製品数を表 2 に示す。表 2 の注意点として、エンドユーザ向けの応用製品では、産業用もしくはオフィスにおける応用製品としており、調査の概要で述べた通り平成 23 年度からコンシューマ向け製品（カーナビゲーション、携帯電話、ゲームなど）を除外している。そのため平成 22 年度から平成 23 年度において、エンドユーザ向け応用製品の件数が大きく減少している。

表 2 分類毎の音声認識製品の件数と推移

分類		表番号*	平成 24 年度		平成 23 年度	平成 22 年度
			件数	増減**	件数	件数
①システム開発者向け 音声認識エンジン	ミドルウェア(組込み向け)	表 B-1	8	0	8	11
	ミドルウェア(PC 向け)	表 B-2	5	-4	9	11
②ソリューション利用者向け音声認識 ソリューション製品	CTI 分野	表 B-3	12	-1	13	10
	Web サービス	表 B-4	3	+2	1	0
	その他	表 B-5	9	+3	6	7
③エンドユーザ向け 応用製品	応用製品	表 B-6	5	0	5	20***
	応用ソフトウェア製品	表 B-7	13	-2	15	21***
合計		-	55	-2	57	80

*表番号：付録 2 製品一覧の対応表番号， **増減：平成 23 年度からの増減

***：平成 22 年度まではコンシューマ製品（カーナビゲーション、携帯電話、ゲーム等）も含む

調査結果の詳細を付録 2（表 B-1～B-7）に掲載する。

本調査結果から、動向を判断することは困難であるが、表 2 からわかるここ数年の傾向を以下に示す。

- ・ システム開発者向けの音声認識ミドルウェアは、特に PC 向けにおいて減少傾向にある。
- ・ ソリューション製品については、CTI 分野はほぼ横ばいだが、Web サービス、その他においては増加傾向にある。特にクラウドコンピューティングを利用したものが増えている。

○平成 24 年度の音声認識・合成技術応用製品の全般的な動向

近年のスマートフォンの普及と、平成 23 年度に登場した Apple 社の iPhone 向け音声対話アプリである Siri、NTT ドコモ社のしゃべってコンシェルが平成 24 年度に国内でも販売されることで、CM 等により一般ユーザに対しても音声関連技術への認知が急速に広まっていった。そのため同様な音声対話アプリの他、音声関連技術を利用したスマートフォン向けの様々なアプリケーションが企画・開発されるようになり、音声認識・合成技術に対するニーズも高まってきた。そのため音声合成製品の動向についてもミドルウェア組み込み向けの製品数が平成 23 年度と比較して大幅に増加していることが見て取れる。一方、サーバ製品は平成 23 年度は一旦増加したものの減少した。またソフトウェアの SaaS またはクラウド化の進展に伴って、サービスについては平成 22 年度から継続して増加傾向にある。音声認識製品の動向についても、クラウドコンピューティングを利用したサービスおよびソリューションが増えつつある。

その他、今回の調査範囲ではないがエンドユーザにおける音声合成技術の応用傾向としては、歌唱音声合成技術である Vocaloid が単なる技術やソフトウェアの範疇を超え、独自性のある 1 つのキャラクターとしてゲーム、音楽、アニメーション等に利用されるようになってきており、さらにこれらの普及は企業ではなく CGM (Consumer Generated Media) 主体で行われている点は、これまで想定されていなかった技術の応用例として特筆されるべき現象であると思われる。

音声認識については、製品動向調査の範囲外の応用として、前述の通りエンドユーザ向けの音声対話アプリケーション、家電の対話的な操作などが増えている。また、会話の共有、もしくはつぶやきの発声を認識して共有することで業務支援を図るアプリケーションの研究開発等、関連技術と有効に連携することで実用を図るものが見られてきた。他には店舗などで音を感知すると、音声認識によってスタンプが貯められるようなスマートフォンのアプリケーションも登場した (JEITA 報告書「音声認識技術の新しい応用に関する調査報告」(2013年6月発行) 参照)。これらの実用化は音声認識技術のモジュール化・低コスト化によってもたらされたものと考えられる。

その他、本調査に含まれない委員会の独自調査からも、スマートフォン、携帯端末のアプリケーションが増加していることが本年度の大きな傾向であり、今後の動向が注目される。

なお、今後も、携帯端末上のアプリケーションや、サーバ上での音声認識・音声合成利用サービスが増加することが予想され、製造企業、販売企業、サービス提供企業などが複雑化するとともに、販売製品という形態ではなく、有償・無償とにかかわらずエンドユーザに対するサービスとして提供される応用システムの増加も考えられる。次年度以降も、なるべく多くの情報を収集し、当該分野の動向を把握するとともに、関連製品開発者、導入検討者、エンドユーザに有益な情報を提供できるよう、調査方法について検討を続ける予定である。