

# クレーン等の音声操作システム

洗 光 範

近年、労働力不足が深刻化する中、建築作業の省人化に貢献すべく、世間一般に普及し始めた音声認識技術と、無線遠隔操作技術を組み合わせ、クレーン等の音声操作システムを開発した。模型実験による音声操作ソフトウェアの検証を経て、実際のジブクレーン、天井クレーン、ホイストおよび遠隔操作散水機に適用し、その有効性と実適用にあたっての問題点を確認できたので紹介する。

キーワード：音声操作、ボイスコントロール、音声認識、信号合図、クレーン、ICT、発話

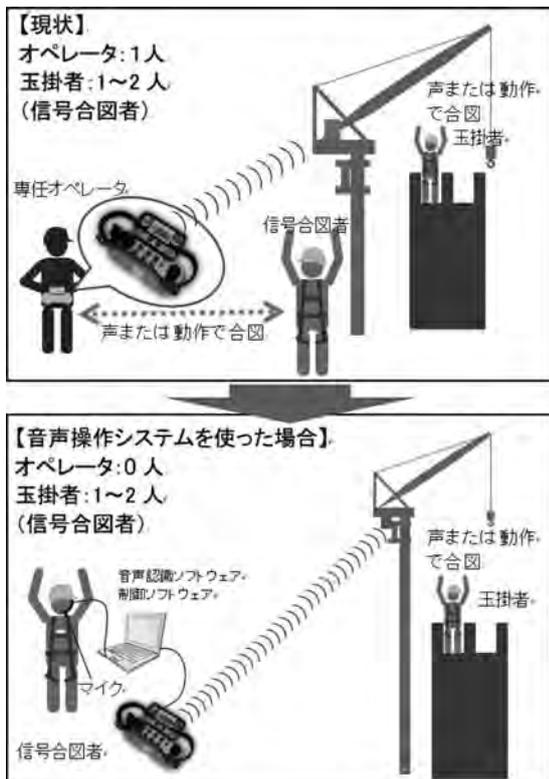
## 1. はじめに

スマートフォンに代表される一般家庭向け IoT 機器の爆発的な普及に伴い、音声認識技術も飛躍的に発達しており、その用途も多岐にわたってきている。一方、国内では人手不足が叫ばれ、我々の建設業界においても、労働力不足が懸念されており、各社とも作業所における技能労働者の確保及び、建築生産性を向上させる技術、機器、方策の開発に取り組んでいる。

クレーンで作業をするには、通常では、玉掛者、信号合図者、オペレータ（クレーン運転者）が必要となる。オペレータは信号合図者の合図通りにクレーンを操作することが、クレーン運転操作の基本であるとされている。また、多くの建築工事現場では、クレーンの信号合図は無線通信機による音声合図が主流となっている。ならば、信号合図者が喋ったとおりにクレーンが動けば、オペレータがわざわざ操作レバーを動かしてクレーンを操作する必要はないのではないか、すなわち必要な人員を一人減らすことができるのではないかと考え、クレーン音声操作システムを開発した(図—1 参照)。

## 2. 音声操作辞書

音声操作の基本技術となる音声認識ソフトウェアとして、市販され、広く世間で利用されているソフトウェアを活用することとした。このソフトウェアの特徴は、アクセントやイントネーションの違いなどに左右されず、認識率が高く、認識した発話が予めユーザが登録した語句と同じであると判断できた場合に、語句に応じたキーコード信号を出力できることである。同時に任意の言葉を発音する機能を兼ね備えているため、実際のオペレータの復唱に近い機能を再現することもでき、音声合図が正しく認識されたか否かが、信号合図者が直ぐに判るため、非常に有効であった。また、音声認識は、無音部分で区切られた1発話単位で行われるため、通常の会話の中に登録された語句が含まれていても信号合図とは認識しないため、誤認識、誤動作の危険性が極めて少ないことも特徴の一つであ



図—1 クレーン信号合図の現状と音声操作

る。これらの機能を活用して、認識する語句と、それに応じた出力キーコードと正しく認識した際に発音する言葉を対応させる表を、ジブクレーン用オリジナル『辞書』ファイルとして作成した。

音声合図の語句は、実際の信号合図者が通常の作業でクレーンオペレータに合図しているのと同じような語句を心がけて選定し、当システムを使うための特別な語句や、喋ってはいけない語句を可能な限りなくし、信号合図者がストレスなく使えるように努めた。例えばクレーン動作の『上げ』、『下げ』を意味する語句として、日本国内で広く使われている、「ゴーへー」、「スラー」を基本として採用した。

音声信号の曖昧さを許容するため、同一のクレーン動作に対して、辞書には複数の語句を登録した。特にサ行で始まる発話は、最初の音を認識しづらいという特性があるため、クレーンの信号合図にとって最も重要な合図である『ストップ』に対しては、「フトップ」、「トップ」などの比較的近い音の語句も辞書に登録し、最初の音を「ス」と正確に認識できなくても動作するようにした。更には、「トマレ」、「テイシ」等の誤認識が生じにくい語句も登録し、認識率を向上させるべく工夫をしている。

### 3. クレーン操作ソフトウェア

今回採用した音声認識ソフトウェアには、認識した語句に応じて、パソコンのキーコードを出力する機能が備わっている。すなわち、パソコンのキーコード番号で制御できる機器であれば、比較的容易に制御することができる。

ジブクレーンの動作は基本的には『起伏』、『巻き』、『旋回』の三動作であるが、これに、速度の増減、三動作のうちの一つのみの指定、チョイ動作（直前の動作を極短時間だけ繰り返す）を行えるソフトウェアを作成した。速度増減の音声合図は、「ハヤク」や「オソク」で現在の速度から1ノッチずつ増減することももちろん、「サンノッチ（3ノッチ）」など、直接動作速度を指定することも可能とした。複数動作作動中に「オヤニソク（親2速）」、「コストップ（子ストップ）」など、一つの動作の動きのみを指定することが可能である。また、前動作をほんの短時間繰り返す『チョイ動作』では、動作時間を機器による応答時間の違いや遅延時間に対応して任意に設定できるようにした。また、先に述べたように、パソコンのキーコードで動作を制御しているので、直接キーボードからキーインしてクレーンを操作することも可能である。

このように、音声合図の曖昧さを許容し、かつ、実際の音声合図に近い発話によるクレーン動作を実現するために、辞書に登録した語句数は約500語、クレーン動作の種類は40近くにもなった。

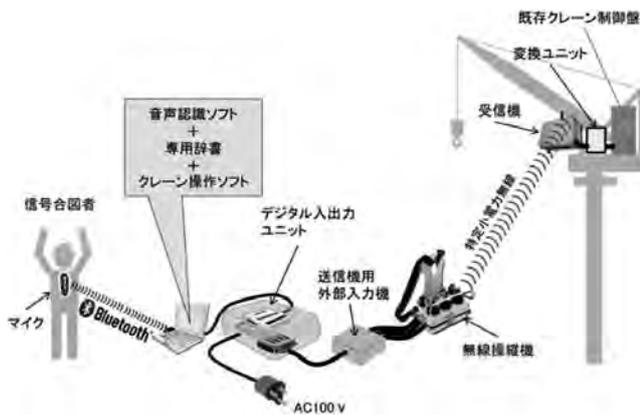
表1に、ジブクレーン用に作成した辞書の一部を示す。表の左列から、動作名称、認識する語句、認識された際に発音される言葉、出力するキーコードを表している。

表1 ジブクレーン用音声合図辞書の例

Name	Spoken	AquesTalk	KeyCode
ジブ起し4	おやおこし	じ'ぶを/おこしま'す。	3/D3
ジブ起し5	おやおこしー	じ'ぶを/おこしま'す。	3/D3
ジブ起し6	おやーげ	じ'ぶを/おこしま'す。	3/D3
ジブ起し7	おやーげー	じ'ぶを/おこしま'す。	3/D3
ジブ起し8	おやおこして	じ'ぶを/おこしま'す。	3/D3
ジブ起し9	おやおこしてー	じ'ぶを/おこしま'す。	3/D3
ジブ起し10	おやごー	じ'ぶを/おこしま'す。	3/D3
ジブ起し11	おこし	じ'ぶを/おこしま'す。	3/D3
ジブ起し12	おこしー	じ'ぶを/おこしま'す。	3/D3
ジブ起し13	おこしてー	じ'ぶを/おこしま'す。	3/D3
ジブ起し14	おこして	じ'ぶを/おこしま'す。	3/D3
ジブ伏せ1	おやすら	じ'ぶを/ふせま'す。	4/D4
ジブ伏せ2	おやすらい	じ'ぶを/ふせま'す。	4/D4
ジブ伏せ3	おやすらー	じ'ぶを/ふせま'す。	4/D4
ジブ伏せ4	おやふせ	じ'ぶを/ふせま'す。	4/D4
ジブ伏せ5	おやふせー	じ'ぶを/ふせま'す。	4/D4
直前動作に対して9	ちょーこーそく	ちょ'ー/こーそく/にし'ます。	0/D0
直前動作に対して10	ちょーこーそくー	ちょ'ー/こーそく/にし'ます。	0/D0
直前動作に対して11	よんそく	ちょ'ー/こーそく/にし'ます。	0/D0
直前動作に対して12	よんのつち	ちょ'ー/こーそく/にし'ます。	0/D0
直前動作に対して13	ちょい	ちょ'い/いきます。	g/G
直前動作に対して14	ち	ちょ'い/いきます。	g/G
直前動作に対して15	ち	ちょ'い/いきます。	g/G
直前動作に対して16	ち	ちょ'い/いきます。	g/G
直前動作に対して17	ち	ちょ'い/いきます。	g/G
直前動作に対して18	ち	ちょ'い/いきます。	g/G
直前動作に対して19	ち	ちょ'い/いきます。	g/G
直前動作に対して20	ち	ちょ'い/いきます。	g/G
直前動作に対して21	ち	ちょ'い/いきます。	g/G
直前動作に対して22	ち	ちょ'い/いきます。	g/G
直前動作に対して23	ち	ちょ'い/いきます。	g/G
直前動作に対して24	ち	ちょ'い/いきます。	g/G
直前動作に対して25	ち	ちょ'い/いきます。	g/G
直前動作に対して26	ち	ちょ'い/いきます。	g/G
直前動作に対して27	ち	ちょ'い/いきます。	g/G
直前動作に対して28	ち	ちょ'い/いきます。	g/G
直前動作に対して29	ち	ちょ'い/いきます。	g/G
直前動作に対して30	ち	ちょ'い/いきます。	g/G
直前動作に対して31	ち	ちょ'い/いきます。	g/G
直前動作に対して32	ち	ちょ'い/いきます。	g/G
直前動作に対して33	ち	ちょ'い/いきます。	g/G
直前動作に対して34	ち	ちょ'い/いきます。	g/G
直前動作に対して35	ち	ちょ'い/いきます。	g/G
直前動作に対して36	ち	ちょ'い/いきます。	g/G
直前動作に対して37	ち	ちょ'い/いきます。	g/G
直前動作に対して38	ち	ちょ'い/いきます。	g/G
直前動作に対して39	ち	ちょ'い/いきます。	g/G
直前動作に対して40	ち	ちょ'い/いきます。	g/G
直前動作に対して41	ち	ちょ'い/いきます。	g/G
直前動作に対して42	ち	ちょ'い/いきます。	g/G
直前動作に対して43	ち	ちょ'い/いきます。	g/G
直前動作に対して44	ち	ちょ'い/いきます。	g/G
直前動作に対して45	ち	ちょ'い/いきます。	g/G
直前動作に対して46	ち	ちょ'い/いきます。	g/G
直前動作に対して47	ち	ちょ'い/いきます。	g/G
直前動作に対して48	ち	ちょ'い/いきます。	g/G
直前動作に対して49	ち	ちょ'い/いきます。	g/G
直前動作に対して50	ち	ちょ'い/いきます。	g/G
直前動作に対して51	ち	ちょ'い/いきます。	g/G
直前動作に対して52	ち	ちょ'い/いきます。	g/G
直前動作に対して53	ち	ちょ'い/いきます。	g/G
直前動作に対して54	ち	ちょ'い/いきます。	g/G
直前動作に対して55	ち	ちょ'い/いきます。	g/G
直前動作に対して56	ち	ちょ'い/いきます。	g/G
直前動作に対して57	ち	ちょ'い/いきます。	g/G
直前動作に対して58	ち	ちょ'い/いきます。	g/G
直前動作に対して59	ち	ちょ'い/いきます。	g/G
直前動作に対して60	ち	ちょ'い/いきます。	g/G
直前動作に対して61	ち	ちょ'い/いきます。	g/G
直前動作に対して62	ち	ちょ'い/いきます。	g/G
直前動作に対して63	ち	ちょ'い/いきます。	g/G
直前動作に対して64	ち	ちょ'い/いきます。	g/G
直前動作に対して65	ち	ちょ'い/いきます。	g/G
直前動作に対して66	ち	ちょ'い/いきます。	g/G
直前動作に対して67	ち	ちょ'い/いきます。	g/G
直前動作に対して68	ち	ちょ'い/いきます。	g/G
直前動作に対して69	ち	ちょ'い/いきます。	g/G
直前動作に対して70	ち	ちょ'い/いきます。	g/G
直前動作に対して71	ち	ちょ'い/いきます。	g/G
直前動作に対して72	ち	ちょ'い/いきます。	g/G
直前動作に対して73	ち	ちょ'い/いきます。	g/G
直前動作に対して74	ち	ちょ'い/いきます。	g/G
直前動作に対して75	ち	ちょ'い/いきます。	g/G
直前動作に対して76	ち	ちょ'い/いきます。	g/G
直前動作に対して77	ち	ちょ'い/いきます。	g/G
直前動作に対して78	ち	ちょ'い/いきます。	g/G
直前動作に対して79	ち	ちょ'い/いきます。	g/G
直前動作に対して80	ち	ちょ'い/いきます。	g/G
直前動作に対して81	ち	ちょ'い/いきます。	g/G
直前動作に対して82	ち	ちょ'い/いきます。	g/G
直前動作に対して83	ち	ちょ'い/いきます。	g/G
直前動作に対して84	ち	ちょ'い/いきます。	g/G
直前動作に対して85	ち	ちょ'い/いきます。	g/G
直前動作に対して86	ち	ちょ'い/いきます。	g/G
直前動作に対して87	ち	ちょ'い/いきます。	g/G
直前動作に対して88	ち	ちょ'い/いきます。	g/G
直前動作に対して89	ち	ちょ'い/いきます。	g/G
直前動作に対して90	ち	ちょ'い/いきます。	g/G
直前動作に対して91	ち	ちょ'い/いきます。	g/G
直前動作に対して92	ち	ちょ'い/いきます。	g/G
直前動作に対して93	ち	ちょ'い/いきます。	g/G
直前動作に対して94	ち	ちょ'い/いきます。	g/G
直前動作に対して95	ち	ちょ'い/いきます。	g/G
直前動作に対して96	ち	ちょ'い/いきます。	g/G
直前動作に対して97	ち	ちょ'い/いきます。	g/G
直前動作に対して98	ち	ちょ'い/いきます。	g/G
直前動作に対して99	ち	ちょ'い/いきます。	g/G
直前動作に対して100	ち	ちょ'い/いきます。	g/G
全ストップ1	すとつぷ	ぜん'どを/い'し'します。	A
全ストップ2	すとつ	ぜん'どを/い'し'します。	A
全ストップ3	すと	ぜん'どを/い'し'します。	A
全ストップ4	すとー	ぜん'どを/い'し'します。	A
全ストップ5	すとーぶ	ぜん'どを/い'し'します。	A
全ストップ6	ふとつぷ	ぜん'どを/い'し'します。	A
全ストップ7	ふとつ	ぜん'どを/い'し'します。	A
全ストップ8	ふと	ぜん'どを/い'し'します。	A
全ストップ9	ふとー	ぜん'どを/い'し'します。	A
全ストップ10	どまれ	ぜん'どを/い'し'します。	A
全ストップ11	とまれー	ぜん'どを/い'し'します。	A
全ストップ12	とめて	ぜん'どを/い'し'します。	A
旋回ストップ1	せん'か'い'すとつぷ	ぜん'どを/い'し'します。	A
旋回ストップ2	せん'か'い'すと	ぜん'どを/い'し'します。	A
旋回ストップ3	せん'か'い'すとー	ぜん'どを/い'し'します。	A
旋回ストップ4	せん'か'い'すとーぶ	ぜん'どを/い'し'します。	A

### 4. 制御機器

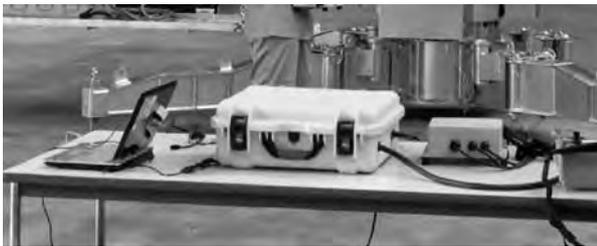
ソフトウェアの機能でキーコードを出力することはできるが、通常のクレーンの操作信号はパソコンのキーコードとは異なるため、キーコード出力に応じて、クレーン操作信号を出力する機器が必要となる。音声認識ソフトウェアとジブクレーン用辞書を組み込んだパソコンと、出力されたキーコードをクレーン操作信号に変換するI/Oコンバータと、無線送信機を



図一 2 ジブクレーン用システム機器構成



写真一 2 実際のジブクレーンへの適用状況



写真一 1 システム機器の例

組み合わせたシステム機器を設計製作した。図一 2 にジブクレーン用のシステム構成図を、写真一 1 に実際に制作したシステム機器の例を示す。

ジブクレーン用システム機器の設計にあたっては、信号合図者からの一方向の通信だけでなく、将来的にはクレーン側からの吊り荷重やジブ角度、各モータの作動状況データなどをフィードバックさせることを想定して、双方向通信が可能な無線送受信装置を選定した。

信号合図者の音声合図を拾うマイクは、高音質で指向性の強いモノを選定し、信号合図者の発話以外の周辺騒音を極力拾わないようにした。

## 5. クレーン実機への適用

音声認識ソフトウェア、ジブクレーン用辞書及びクレーン操作ソフトウェアの動作を、パソコンの出力信号で動作するクレーンの模型で検証した後、前章で述べた制御機器を製作し、実際のクレーンに適用した(写真一 2)。

図一 2 に示したように、クレーン側(受信側)には、無線電波の受信機及び変換ユニットを設置し、受信機で受信した信号をクレーンの制御電気信号に変換して、クレーンを操作する。一方、信号合図者側(送信側)のシステムは、複数の必要機器が有線でつながられ、またパソコンやデジタル入出力ユニットを作動させ

るための電源ケーブルが必要となる。信号合図者は荷取り位置や、荷卸し位置に応じてある程度移動する必要があるが、これらの機器と共に自由に移動することは困難である。従って、今現在では、システム機器を合図者の近傍に設置し、無線マイクの電波(今回はBluetooth)が繋がる範囲内での移動に限定した。

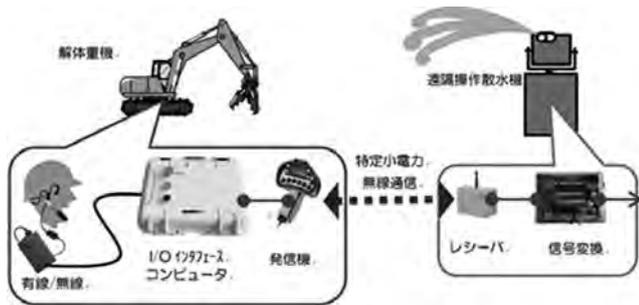
現在は法律上の規制もあり、我々が勝手に無線機器を改造したり部品を転用したりすることはできないが、今後の電子機器および無線通信技術の発達に伴って、手軽に身に付けて持ち運べるような小型化が実現することを期待する。

## 6. ジブクレーン以外の機器への適用

先にも述べたが、当該システムは無線操作が可能な機器であれば、ほとんどの機器に適用可能である。クレーンの場合には専任のクレーン操作者が不要になるといったメリットがあったが、他の機器に応用する場合、人員削減のほかにもいくつかのメリットとデメリットが考えられるので、その適用例をいくつか紹介する。

### (1) 遠隔操作散水機への適用

今後需要が高まることが予想される、老朽ビル等の破碎解体工事において、周辺環境への配慮から粉塵の飛散防止のために散水養生を行っている。一般的な散水作業には、ハイウォッシュャや高圧洗浄機等を使用しており、破碎箇所への移動に伴い散水場所も移動する必要がある。そのために散水専用の作業員を配置しているのが現状である。しかし、破碎箇所の近傍での散水作業は、破碎ガラの飛来や重機との接触など、被災リスクの高い作業である。そこで、最近では散水機の放水銃の向きを、離れた場所からリモコン等で操作して、被災リスクを低くした製品も普及し始めている。



図一三 遠隔操作散水機用のシステム機器概要



写真一三 遠隔操作散水機の実機実験

この遠隔操作散水機の散水方向(放水銃の向き)を、解体用重機のオペレータが自分の発話で直接指示することによって、散水専用の作業員を減らすことが可能となった。図一三に遠隔操作散水機のシステム機器概要を、実験状況を写真一三に示す。

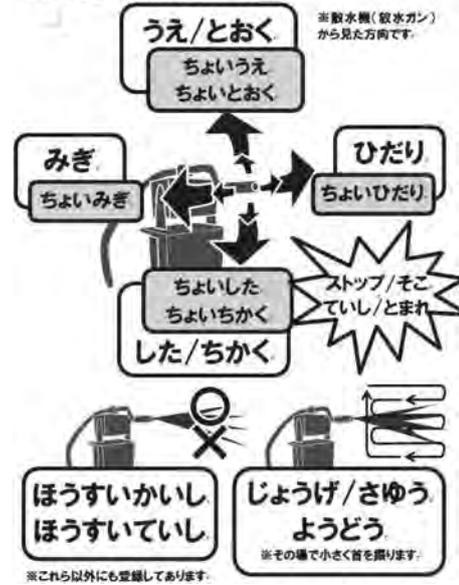
また、クレーンと同様に操作自体はパソコンで制御できるため、付加価値として、自動首振り機能を追加した。音声合図「ヨウドウ」を認識すると、その場で小さく左右に首を振りながら上下方向にも移動し、比較的広い範囲に散水できる様にした。首振り方向を左右のみ、上下のみに限定するなど、複数種類の自動首振り機能も装備した。図一四に遠隔操作散水機用の音声合図の例を示す。

実際のビル解体現場では、当然のことながら大きな騒音があり、音声認識率が低下してしまった。現場状況に合わせてマイク感度の調整を行う必要が生じたり、騒音環境下でも認識されやすい喋り方を修得する必要が生じたり、新たな課題も見えてきた。

(2) 天井クレーンへの適用

工場などの恒久施設で多く使われている天井クレーンや橋形クレーンでは、玉掛者が荷を玉掛けした後、片手でクレーンの操作スイッチを操作しながら、もう一方の手で荷を介錯するといった場面が多く見受けられる。このような場面では、操作スイッチを使わずに

主な音声操作合図例



図一四 遠隔操作散水機用の音声合図例

音声合図だけでクレーン操作ができれば、両手で荷の介錯が行えるので、より安全な作業となる。

天井クレーン用のシステムは、ジブクレーン用をベースにして、より簡易なものとした。パソコンおよびモニタを小型なものに変更し、I/Oコンバータなどインターフェース機器と共に一つのボックス内に収めて、コンパクトにした。また、バックアップバッテリーを追加して、電源ケーブルを抜き差しするだけで、パソコンの起動と遮断、及び各機器の電源の入切を自動で行えるようにし、システムの起動時や停止時の面倒な手順をなくして、ユーザにとってより使い勝手の良いシステムとした。送信機側システム機器を写真一四に、実際の天井クレーンへの適用状況を写真一五に示す。

送信側機器は天井クレーンの走行範囲の中間付近の見通しの良い位置に設置し、オペレータは無線マイクのみを身に付けて音声合図で操作している。オペレー



写真一四 コンパクトな天井クレーン用送信側機器



写真—5 天井クレーンへの適用状況

タと送信側機器までの距離は、周辺環境やマイク性能にもよるが、この例では 30 m 程度でも確実に操作できていることを確認できた。

### (3) ホイストへの適用

ALC 工事では重たい部材を、ホイストを用いて建て起こし、吊り上げて取り付けている。しかし、ホイストの設置階と ALC の取り付け階が異なることも多く、ホイストの操作のために、専用の作業員を配置していることが多い。そこで、ホイスト操作を音声合図で行えるようにすることで、専用のホイスト操作者を配置することなく、部材の取り付け者が直接ホイストを操作できるようにした。

しかし、ホイストは、『上げ』『下げ』の 2 操作しかなく、かつ吊り上げ距離も短く、作業員もすぐそばで作業しているため、他のクレーン等に比較してより早いレスポンスを求められた。音声認識は、話者の発話を認識すると共に、発音の空白を区切りとして認識して、その発話単位で登録された辞書内の語句との照合を行っている。したがって、時間遅れが必ず生じてしまう。これが、大きなクレーンや散水機などでは気にならないレベルであったが、より早いレスポンスを求められるホイスト作業にあっては不評で、実機での適用には至らなかった。

## 7. おわりに

ICT の発展に伴い、音声認識、音声操作技術も様々な生活の場面で活用されるようになってきた。しかしながら、音声合図で実際の物理的な物が動くことはほとんど例を見ない。人間同士の会話でも、聞き取りにくいことはあり、聞き直すこともあるだろうし、正しく認識できずに間違った反応をして、指示を出した者に怒られることもあるであろう。音声合図は 100% 正確に伝わるわけではないため、安全にかかわる行為に採用するには、必ずリスクを伴う。安全を考慮しすぎるあまり、音声合図を認識できなかった時には動作を止めるなど、フェールセーフを強化しすぎると、実際には非常に使い勝手の悪いシステムとなってしまう、実用に耐えられなくなるおそれが出てくる。不安全リスクを許容することはできないが、如何にして操作性を高めていくかが大きな課題となっている。

また、冒頭にも述べたが、「クレーン運転者は合図者の合図通りにクレーンを操作する」ことを基本に当該システムの開発をはじめた。音声合図の調査のために実際のタワークレーンのオペレータにヒアリングを行った中で、「オペレータは必ずしも合図者の合図通りに操作しているわけではない」「(合図者の喋った)合図を復唱しながら合図と違う操作をやることもある」「(ジブ先端から下を映す)カメラ(映像)がないと怖くて運転できない」といった声を聞くことができた。実際のクレーンでの揚重作業には、オペレータの技量、知識、思考、判断と合図者との信頼関係によって安全で確実な作業が行われていることを再確認した。

音声認識機能を利用したモノの操作や制御には、いまだリスクが伴うが、同時に応用範囲が急激に広がっている AI 技術等を併用することによって、先に紹介したようなオペレータの技量や判断を再現し、更に使い勝手の良い、より安全性を高めたシステムを実現することができると考えている。音声認識技術や音声操作技術の今後の更なる発展に大いに期待している。

JICMA

#### [筆者紹介]

洗 光範 (あらい みつり)  
 榎竹中工務店  
 生産本部 生産企画部  
 部長 機械電気担当

