マイクロホンアレイ　可視光通信　指向性

光で音をつかまえる――LEDと高速カメラで挑む超多チャネル音響信号の観測

NTTコミュニケーション科学基礎研究所では大規模な空間の音響信号をまるごとつかまえるため，並列性に優れた高速カメラとマイクロホン信号をLEDに変換するセンサからなる可視光通信を利用し，マイクロホンアレイを試作しました．そして配線や電波に制約されずに，120チャネルのマイクロホンアレイ信号による実時間の指向性収音を実現し，将来の大規模化の見通しを得ました．これによりライブ・イベントなどでの超指向性収音や臨場感音響環境の実時間録音，取得・配信への応用が期待されます．

Pablo Nava Gabriel /白木（しらき） 善史（よしふみ）
Nguyen Hoang Duy /鎌本（かまもと） 優（ゆたか）
佐藤（さとう） 尚（たかし） /原田（はらだ） 登（のぼる） /守谷（もりや） 健弘（たけひろ）
NTTコミュニケーション科学基礎研究所

はじめに

複数のマイクロホンを並べるマイクロホンアレイは，マイクロホンの数を増やすことで，指向性収音，音響空間を丸ごと観測し，臨場感再現，雑音除去などの機能を高めることが可能です．現在までに，マイクロホンアレイによる指向性収音[(1)]，臨場感再現のための波面合成[(2)]などはおよそ100チャネル程度までは数多くの実験例があります．一方で，例えば2 m四方の部屋において4 kHzまでの音波の状態をすべて収集するためには，約2500個のマイクロホンが必要になります．これを拡大し，コンサートホールや競技場の音響空間を再現しようとすれば膨大な数のマイクロホンを使用して録音しなければなりません．

また指向性についてもその性能を高めるには，やはりマイクロホンの数を増やすことが必要です．マイクロホンからアンプやAD変換器までの比較的長い配線，AD変換器の同期などを考えると1000チャネル，1万チャネルへの拡張，大きな空間での配置などの実現には高いコストや難しさが予想されます．さらに電波でも膨大な数の音響データを同期して伝送するには多くの電波帯域を消費しますので，無線電波であっても多チャネルの同期した情報伝送の実現は困難を伴います．

全体構成

前述のような困難が予想される一方で，近年LEDや光検出器およびカメラの高性能化に伴い，可視光通信の研究が進展しています．中でも多数の送信機からのデータを並列的に取得する目的には高速カメラを使用する可視光通信が有用です．特に近年の速度と画素数の増大の進歩は著しく，コストダウンが図られており，画素数を制約すれば1秒に最大10万フレームもの撮影が可能な市販品の高速カメラを入手することができます．さらに画像データのような多くのデータの並列処理にはGPGPU（General Purpose Graphic Processing Unit）＊1が適しています．GPUは本来画像描画処理のための高速並列プロセッサですが，より一般的な並列信号処理にも適用可能です．以上のような背景から，NTTコミュニケーション科学基礎研究所では**図1**

＊ GPGPU：一般の高速並列計算目的で使うグラフィックス描画用並列演算器．

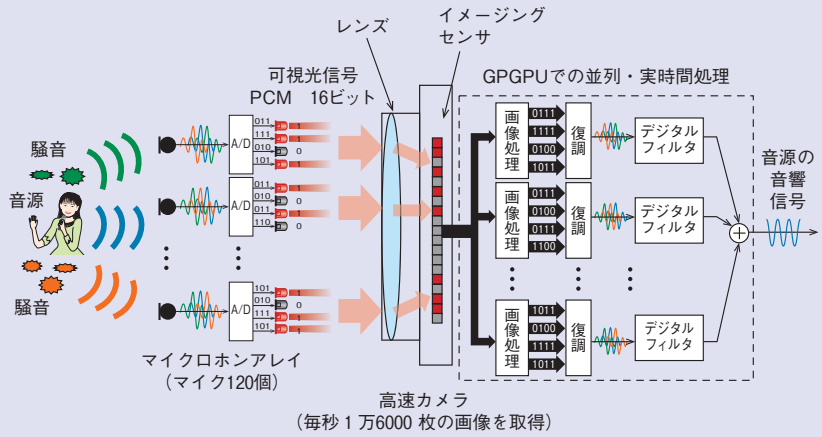


図1　システム構成

のようなLED送信機を搭載する音響センサと受信用の高速カメラ，GPGPUを使った高速並列画像処理を行うシステムを試作し，配線や電波帯域に制約されない，自由度の高いマイクロホンアレイで実時間の指向性収音を実現しました[3]．

音響センサ

音響センサはマイクロホンとAD変換とLED，さらに制御用の赤外線受光部からなり，電池または外部からの5 V電源の入力を必要とします（**図2**）．音響信号はデジタルマイクによって取得し，これを16 kHzのリニアPCM16ビットの信号に変換します．このビット情報を4×4個，計16個のLEDの0，1に対応させた点滅によって送信しています．音響センサの配置は自由ですが，今回の実験では4 cm間隔とし，縦5個×横40個の計200個を配置しました（**図3**）．なお，送信機には可視光LEDのみならず，目に見えない赤外線LEDを使用することも可能です．音響センサは自走用の基本クロックを持ちますが，単体では長期にわたり各センサとカメラの同期を維持するのは難しいため，マスタークロックのパルス列を**図4**の赤外線LEDで各センサに分配し，**図5**の高速カメラの駆動にも同じパルス列を用いることで，音響センサのLEDの発光タイミングと高速カメラのシャッタータイミングと同期させています．

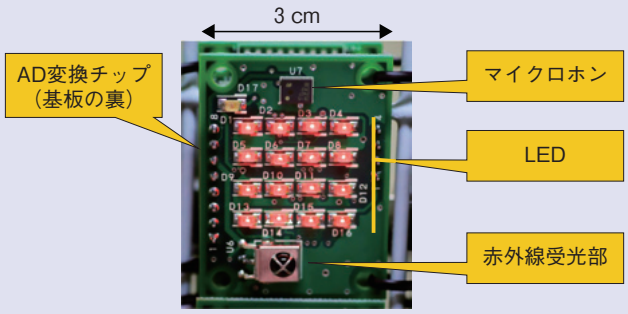


図2　音響センサ

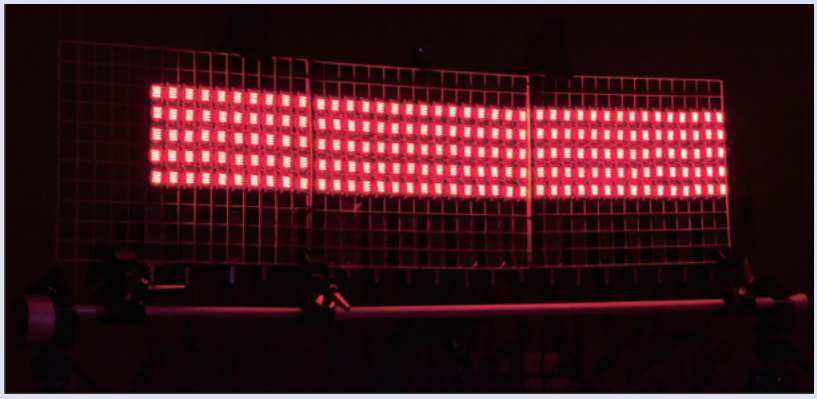


図3　音響センサのアレイ

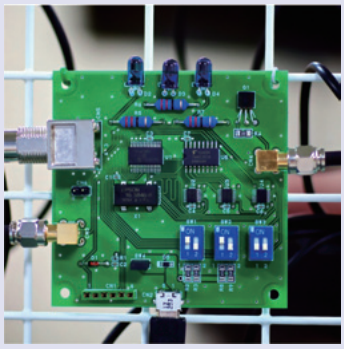


図4　同期用赤外線LED

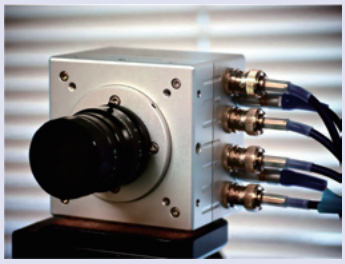


図5　高速カメラ

高速カメラ

画素数は縦100×横800で，200個の音響センサのうち，縦4個×横30個，計120個の情報を音響センサと同じ1秒に1万6000フレームで画像を取得しました．**図6**は実際に取得したLEDの画像例です．なお，通常のビデオカメラでは1秒間に30，ないし60フレームが一般的です．実験ではカメラのレンズを変えることにより30 m程度の距離においても通信可能であることを確認しました．

画像処理と音響処理

取得した画像データ（図6）はGPGPUとの共有メモリに転送し，画像から音声信号を再構成します．そのために，画像認識の複数アルゴリズム

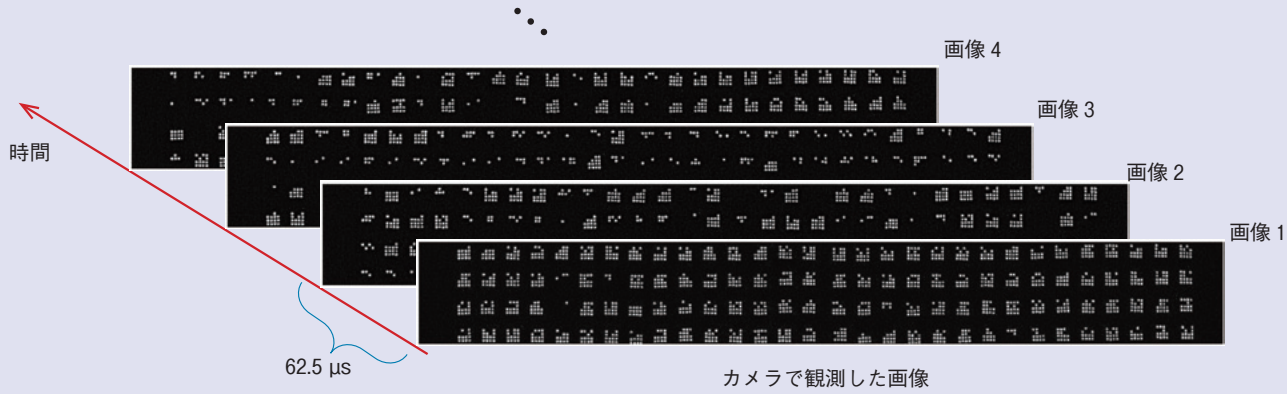


図 6　高速カメラで取得したLEDの画像例

を使い[4]，画像中におけるそれぞれのLEDの位置を推定します．次にそれぞれのLEDの位置にあたるピクセルの輝度解析を行い，ピクセルの輝度信号からそれぞれのビット情報を復号します．１つのLEDについての点滅を複数の画素に対応させて検出する際に，時間方向で画素信号を最大256点蓄積して隣接する画素との識別精度を高めることで，LEDの点滅を誤りなく検出できています．

画像から復号した多チャネルの信号の位相を制御することで指向性を持たせることができます．今回の実験では音質の劣化が少ないといわれている遅延和処理を行いました．これは平面波の到来遅延を仮定し，各チャネルの信号の遅延和処理を行い，目的の方向の信号を強調することができます．また，受信側のユーザが聞きたい音源の方向を設定し，実時間でマイクロホンアレイの指向性の向きを変えることも可能です[3]．

実験では，実時間で指定した方向からの音が強調されて聞こえることが確認され，特にマイクロホンアレイの正面方向からは10 m先からの通常の会話でも聞き取ることができました．

今後の展開

現在のGPGPUでも１万チャネル以上の実時間処理が可能です．加えて進歩の著しいカメラやLEDなどの性能向上を考慮すれば，近い将来，より大規模な音響計測システムの実現が可能になる見込みです．これにより頑健性の高い指向性収音や超臨場感音場伝送などへの応用が期待されます．また，マイクロホンの位置とマイクロホンからの音をカメラで同時に取得できること，マイクロホンの位置が動いても追跡できることなどの利点を活かした応用展開が考えられます．さらに，映像と音声を組み合わせた情報利用などへの展開も期待できます．

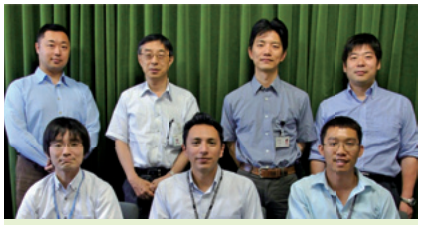
■参考文献

(1) K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda：“Sharp directive beamforming using microphone array and planar reflector,” Acoustical Science and Technology, Vol.34, No.4, pp.253-262, 2013.

(2) S. Koyama, Y. Hiwasaki, K. Furuya, and Y. Haneda：“INVERSE WAVE PROPAGATION FOR REPRODUCING VIRTUAL SOURCES IN FRONT OF LOUDSPEAKER ARRAY,” Proc. EUSIPCO 2011, pp.1322–1326, Barcelona, Spain, Sept. 2011.

(3) G. Pablo Nava, Y. Kamamoto, T. G. Sato, Y. Shiraki, N. Harada, and T. Moriya: “Simultaneous acquisition of massive number of audio channels through optical means,” In Proc. 135th AES Convention, p.8965, New York, U.S.A., Oct. 2013.

(4) G. Pablo Nava, Y. Kamamoto, T.G. Sato, Y. Shiraki, N. Harada, and T. Moriya: “Image processing techniques for high speed camera-based free-field optical communication,” IEEE Int. Conf. ICSIPA, pp.384-389, Melaka, Malaysia, Oct. 2013.



（後列左から）鎌本　優/ 守谷　健弘/
佐藤　尚/ 原田　登
（前列左から）白木　善史/
Pablo Nava Gabriel/
Nguyen Hoang Duy

ここで紹介したのはこれからの研究の入り口であって，LED，高速カメラ，GPUの発展や強みをうまく利用して新しい研究分野の開拓やこれまでには不可能と思われた新しい応用に挑戦していきたいと思います．

◆問い合わせ先
NTTコミュニケーション科学基礎研究所
守谷特別研究室
TEL　046-240-3141
FAX　046-240-3145
E-mail　moriya.takehiro@lab.ntt.co.jp