

# 拍手音を用いた対話的イルミネーション演出

宇根健一郎†1 辻春樹†1 田口博史†1 栗山繁†1

**概要:** 公共の場でのイルミネーションは、その場の雰囲気演出のみならず、地域活性化の効果も期待できる。本研究では、豊橋市が主催する駅前地区景観のイルミネーション化事業で開発された、見学者参加型のイルミネーション演出システムを紹介する。このシステムでは特別な道具を使わずに誰もが簡単にイルミネーションを操作できる仕組みを提供するために、拍手音を用いた対話的な操作方法を開発した。本報告では本イルミネーションの画像投影等に基づく調光演出機構と、対話的な操作手段を提供するための拍手音の識別手法と音高による分類手法を提案する。

## Interactive illumination with clapping sounds

KENICHIRO UNE†1 HARUKI TSUJI†1  
HIROSHI TAGUCHI †1 SHIGERU KURIYAMA†1

### 1. はじめに

本研究では、イルミネーションのデジタル調光と画像処理技術を融合させた、拡張性を有する対話的な演出が可能 LED イルミネーション装置を提案する。これは、豊橋市の駅前商店街の活性化を目論んで、産学官の共同体で開発したプロジェクトの成果である。

デジタル調光可能なカラーLED光源が大量生産されるに伴い、一部のテーマパークでは一千万球を超える大規模なイルミネーション施設が建造されている。しかしながら、本プロジェクトによるイルミネーション展示は営利目的ではなく限られた予算でシステムを構築する必要があるため、小規模でも人を惹きつけられるエンターテインメント性を有する演出効果を検討した。また、製品化も視野に入れた事業であるため、利用シーンに適した演出を専門的な知識を要せずに容易にカスタマイズできる仕組みを構築した。

### 2. イルミネーション装置の概要

#### 2.1 ハードウェア構成と制御機構

イルミネーションツリーは高さ 2.4m の円錐型の支柱に直列に配置された LED 光源を約千個巻きつけて制作した。このリボン状の LED は光源球の配置間隔が 10cm であり、上下の配置間隔が約 4cm となる様に螺旋状に巻きつけた。図 1 にツリーの一部を拡大した写真を示す。

各 LED 光源は、RGB の各チャンネルで 32 階調の調光が可能なので約 3 万色を表現でき、その制御にはリアル・ペリフェラル・インタフェース(SPI) のプロトコルが用いられている。しかし、本システムでは舞台用照明等の分野で演出効果に広く用いられている DMX512 プロトコルを用い、SPI 通信用のデコーダーを用いることで調光制御の信号を各 LED 光源に送出している。

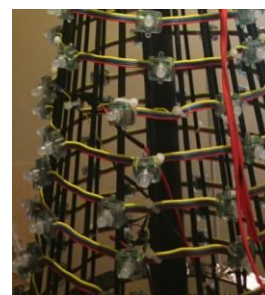


図 1. リボン型 LED の支柱への巻きつけ

シリアル通信の一種である DMX512 は、一つの出力ポートで 8 ビットの信号を最大 512ch まで、1 秒間に 44 回送信できる。今回使用した製品では各 LED 光源に対して RGB の 3 チャンネルを割り当てるので、全体の制御には少なくとも 7 つの DMX 出力ポートが必要となる。このような多数の DMX 出力をコンピューターから扱うには高価な機材が必要であり、ソフトウェア的にも煩雑である。そこで本システムでは、大規模な舞台照明等で使われている、複数の DMX 信号を 1 つにまとめて制御できる調光用のプロトコルを用いて PC から信号を送信する。このプロトコルは Art-Net と呼ばれ、UDP (User Datagram Protocol) の一種であるので、一般的なコンピューターでも容易に出力可能であり、ソフトウェア的にも実装しやすい。さらに、1 本の Ethernet ケーブルで全 LED を制御できるので、システムをコンパクトに構築できる。

図 2 に、上述した制御信号送出のための機器構成を示す。

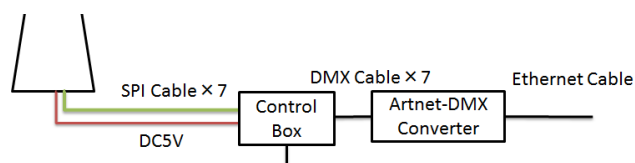


図 2. イルミネーション制御信号の送出機器構成

†1 豊橋技術科学大学 情報・知能工学系  
Toyoashi University of Technology.

## 2.2 遠隔からのネットワーク・メンテナンスシステム

ツリー状のイルミネーション装置(以後、ツリーと呼ぶ)は、大学から遠く離れた豊橋駅前に設置された。ゆえに、2ヶ月半にわたる長期間、屋外で安定に稼働させるために、ソフトウェアやハードウェアの障害が発生した際には遠隔から復旧できる機構を構築するために、VPN接続とvPro技術を用いてメンテナンス用のネットワークを構築した。

本システムではツリーをEthernet上で制御しているので、メンテナンス用のネットワークも制御用のものを拡張してVPNサーバーに遠隔地から接続できるようにした。

Intel社が開発したvProは専用チップセットが搭載されているPCを必要とするが、ネットワーク経由での画面表示、起動、および強制終了等が遠隔操作できる。リモート接続したPCの画面をソフトウェア的に表示するアプリケーションは数多く存在するが、vProはハードウェアベースの技術なのでBIOSレベルの設定が可能であり、ブルースクリーンエラー等の症状が発生する障害時においてもレポート等による復旧が可能である。

屋外の公共スペースでインターネットに接続するために、今回は無線LANを用いてグローバルIPアドレスを取得できるようにWimax2で利用した。ただし、取得が可能なIPアドレスが固定されていないので、ダイナミックDNSを用いた。また、今回は2つのツリーの筐体を物理的に離して設置したので有線での接続が困難となったので、図3の様に無線APを用いてツリー間を接続した。ただし、制御用のPCとツリー間に関しては、高速で安定した経路が求められるので有線で接続した。

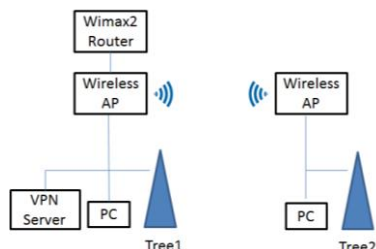


図3. ツリーのネットワーク構成

## 3. 調光演出用アプリケーション

今回製作したイルミネーション装置は、リボン状のLED光源群を図4の様に螺旋状に巻きつけたので、個々のLED光源は横方向にはほぼ等間隔に配置されるものの、周回毎に半径の異なる円錐形状に巻きつけられるので、縦方向には不揃いな配置となる。ゆえに、光源の実空間での位置と2次元格子状に配置された画素の位置の対応関係を計算し、画素の色情報から光源の点灯色を最適に決定する必要がある。本研究では、イルミネーションの各光源色を規則的に自動生成する機構と、表示させたい絵柄図形の画像情報から光源の点灯状態を計算するソフトウェアを、以下のモジ

ュール群から構成した。

- 各LED光源の3次元位置を推定する機能
- 円錐の展開図と絵柄図形を2次元座標上に同時投影し、画素値から光源の色情報をサンプリングする機能
- 各光源の時間遷移状態の規則に基づいて点灯色を計算する機能
- 絵柄や規則的点灯パターンの再生順序や諸変数の値をテキストファイルとして入出力する機能
- 光源の色の値をDMX信号に変換し、イルミネーションの調光制御部に送信する機能



図4. ツリー状イルミネーション装置の外観

調光の演出設計を支援する既存のアプリケーションは、事前登録されたパターンを選択して編集する手段や、画像からパターンを生成する手段等を提供している。我々の開発したシステムでは、演出パターンの選択と画像の投影という双方の特徴を組み込んだハイブリッド型として位置付けられるが、前者の演出パターンに関してはセル・オートマトンやパーティクル・システムに基づく複雑なパターン生成も可能とし、後者の画像投影に関してはシークエンスの設計や非平面上の光源の配置にも対応できる様にした。

2次元のディスプレイ空間で表示される絵柄を3次元のイルミネーション光源の点灯に対応づける投影写像(イルミネーション・マッピングと呼んでいる)の結果は、図5の様に3次元CGライブラリを用いたシミュレーション可視化ツールにより表示され、絵柄の大きさや配置を円錐の展開図上で対話的に編集するツールも実装した。

登録された複数の演出パターンは、後述する拍手音により切り替えられるが、拡張性を持たせるためにその規則はJSON (JavaScript Object Notation) 形式のテキストファイルとして記述できる様にファイル入出力機構を実装した。

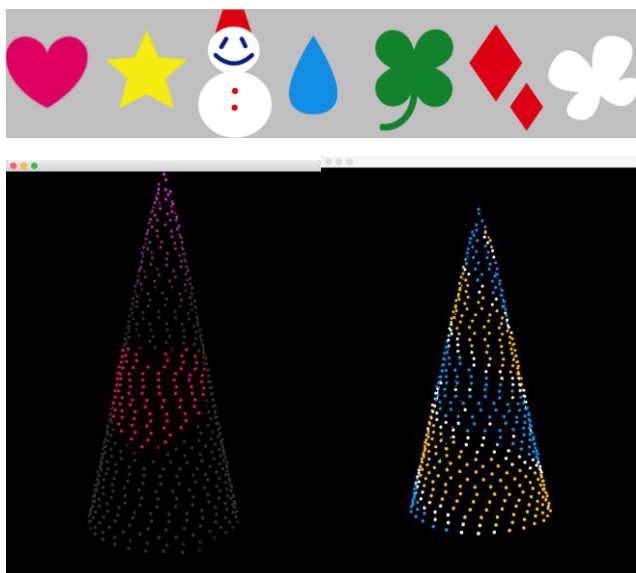


図5. 円錐状のツリーに投影した絵柄図形の標本例と可視化ツールによる点灯状態のシミュレーション例

#### 4. 手を叩く音の検出機構

既存のイルミネーション施設では、その娯楽性を高めるための仕掛けとして、対話的な操作環境の導入事例が既に存在する。例えば、スマートフォンのアプリを用いて点灯パターンを操作するシステム等が既に開発されているが、今回の様な駅前での展示では、帰宅を急ぐ行人でも手軽に参加できる仕組みが望ましいと判断した。したがって、スマートフォン等の機器や特殊な道具は一切使用せずに、身体を用いた対話のみで操作が可能な方法を検討した結果、手を叩く音の特徴によって点灯パターンや絵柄を切り替える方式とした。さらに、「意図的に制御できる楽しさ」と同時に「偶発性による面白さ」も体験できる様に、製作した2台の装置の片方は個人だけの拍手音、他方は2人以上の同時の拍手音に反応する仕組みを導入し、その音の高さ（2人以上の場合は、その高さの組み合わせ）に対して特定の絵柄が選択される機構を開発した。

技術的には、手を叩く音（以後、拍手音と呼ぶ）を他の音から分離する必要があるが、今回は実測した豊橋駅前の環境音の成分を除去しながら拍手音の特徴を表す周波数帯の強度を参照して分離する機構を開発した。以下、拍手音の検出および識別の方法について述べる。

##### 4.1 拍手音の検出

拍手音の波形は、図6の様に約0.01秒ほどの短く大きいインパルスが現れた後に振幅が減衰し、約0.1秒で消失する様な振幅形状をしている。また、図7より10kHz以上の高周波数帯にも複数のスペクトルピークが確認できるが、その出現頻度や周波数は個人間の変動が大きいことが判明した。これは、拍手音のスペクトルが手のひら同士をぶつ

ける角度やタイミングで敏感に変化することに起因すると考えられる。ゆえに、拍手音は図6に示すインパルス音を検出することとした。

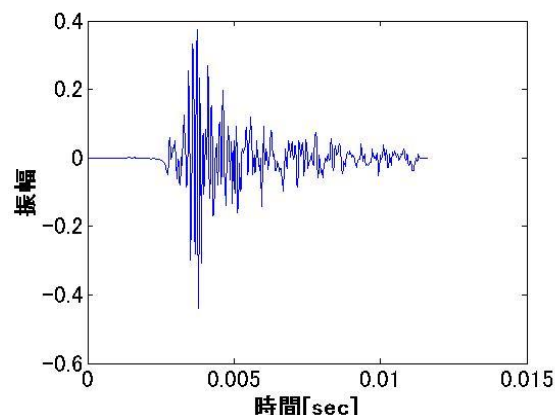


図6. 拍手音の振幅の時間変化

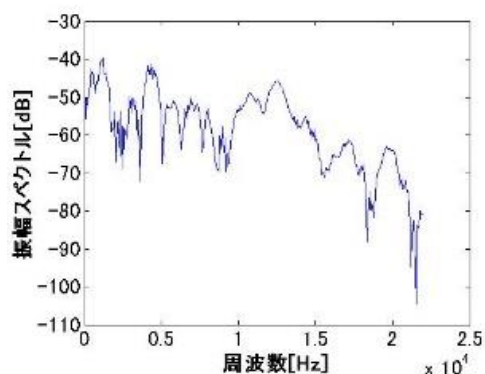


図7. 拍手音の振幅スペクトルの周波数分布

実際の拍手音を録音して解析した結果、拍手音の多くは低周波の減衰が遅く、高周波の減衰は速い傾向が確認された。ここで、低周波の音は空気による吸収の影響が少ないことが知られている[1]。このことから、豊橋駅前で観測される残響音は壁などの障害物による反射の影響を受けず、空気吸収の影響のみを受けていると考えられる。ゆえに、拍手音を安定に検出するため、残響音の特徴も追加した。残響音特徴は各周波数での音の減衰により求められるが、拍手音の残響はインパルスの発生から0.1秒ほどで消失するので、短時間フーリエ変換(Short-Time Fourier Transform:STFT)を用いて解析した[2]。これは、図8に示す様にSTFTで得られた時間-周波数解析結果から、各周波数成分での時間経過による音の減衰を直線近似した際の勾配値から求められる。ただし、STFTは波形を切り出すフレームを移動させながら周波数解析するので、フレーム番号軸は時間軸に相当することに注意されたい。

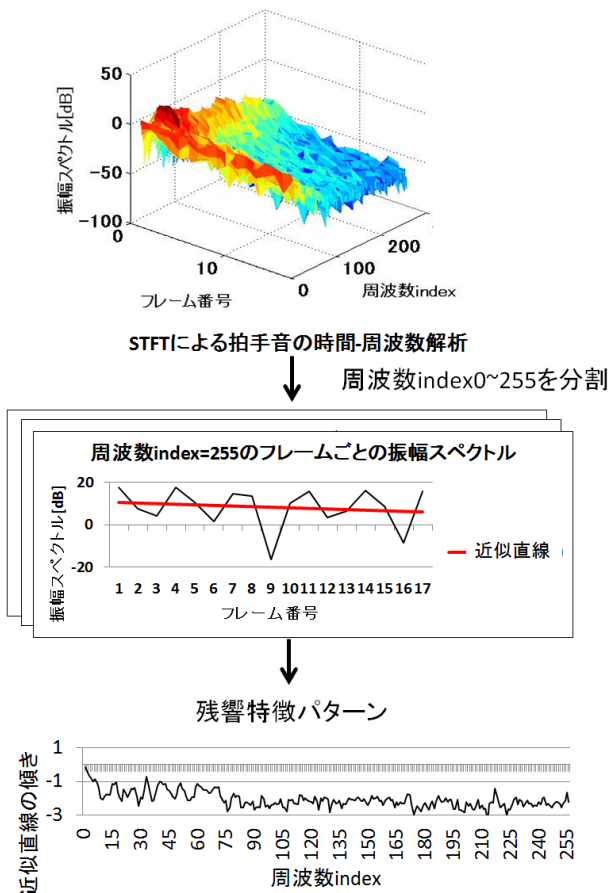


図8. 残響特徴パターンの算出法

以上の特徴を考慮して、図9の拍手音検出アルゴリズムを提案する。

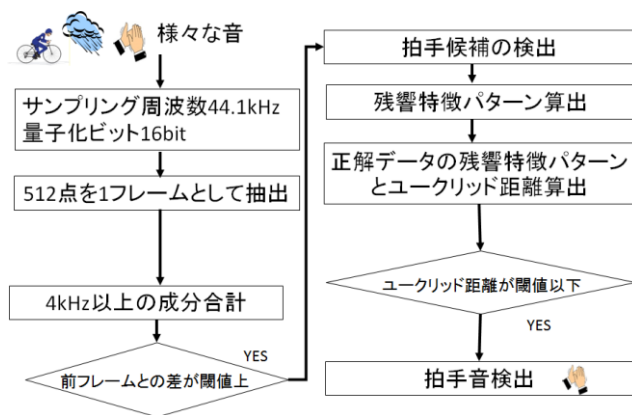


図9. 提案する拍手音検出法

まず、式(1)により求めた前フレームとの差  $d_T$  が閾値を超えたとき、その音を拍手音候補として検出する。

$$d_T = \sum_{k=47}^{N/2-1} |X_k^T| - \sum_{k=47}^{N/2-1} |X_k^{T-1}| \quad (1)$$

ここで、 $N$  はフレーム長、 $T$  はフレーム番号、 $k$  は周波数のインデックスである。また、 $|X_k|$  は時間信号  $x(t)$  のフーリエスペクトルを  $X_k$  とした時の振幅スペクトルである。式

(1) では、4kHz 以上の振幅スペクトルの差を求めるため、周波数インデックス  $k$  が 47 以上の振幅スペクトルを対象としている。

この段階では、音量が大きいものを全て検出してしまうので、検出された音  $T$  に対して残響特徴パターンを算出し、式(2)で求めた正解データ  $R$  とのユークリッド距離  $d(R, T)$  が閾値以下の場合に、拍手音として検出する。

$$d(R, T) = \sqrt{\sum_{i=0}^{N/2-1} (r_i - t_i)^2} \quad (2)$$

## 4.2 拍手音の音高分類

拍手音の周波数特徴量は、図7に示す様にスペクトルに変動が認められるものの、このスペクトル変動の個人の特徴や、全ての拍手音に共通した特徴を捉えることができれば、その種別の分類に利用できると考えられる。

ここでは、拍手音のスペクトルピークが一定の周波数帯で変動していると仮定し、5414個のサンプルに対して4kHz以上の最大スペクトルピークを算出した。ここで4kHz以上と制限したのは、豊橋駅前の環境音が4kHz以下に集中していることが事前調査により明らかになったためである。このスペクトル特徴から、環境音の影響を受けない特徴量を抽出すると、スペクトルピークの最大値は約6.8kHz、最小値は約4kHz、平均は約4.4kHz、および標準偏差は約320Hzとなった。ここで、スペクトル変動が正規分布に従うと仮定すると、99.7%の拍手は、3480Hz~5320Hzの周波数帯にスペクトルピークが現れると考えられる。ゆえに、拍手音のスペクトル最大値は1840Hzの変動があると仮定し、これを拍手音の音高の特徴とする。

拍手音の音高特徴に対する認識には個人差があり、音高が異なる様に手を叩くように指示しても、観測される波形データに一貫した特徴は見られなかった。そこで今回は、個人間の変動が小さい3種の音高の拍手音を得るために、図10に示す3つの叩き方を指定して録音したデータを分析に用いた。



図10. 実験で指定した3種類の手の叩き方

拍手音が1840Hzの変動があると仮定し、周波数の変動を吸収した距離の算出のために、2つの特徴パターンを非線形にマッチングするDPマッチングを導入した。ただし、雑音の影響を減少させるためケプストラム分析によるスペクトル包絡に対してDPマッチングを適用した[3][4]。提案する拍手音の音高分類手法の処理の流れを図11に示す。

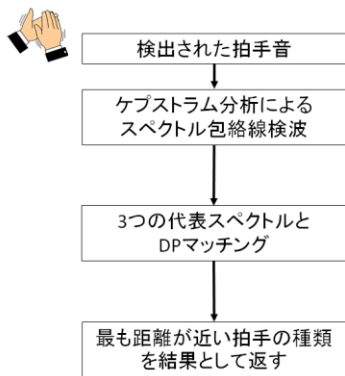


図 1 1. 提案する拍手音の音高分類の流れ

### 4.3 拍手音の検出および識別の検証

約3時間のwavデータに対して拍手の検出精度を検証した結果、実際に得られた946回に対して正しく区間を検出した割合である検出率は92%、拍手と検出した区間が実際の拍手と一致した(誤った検出を含まない)割合の適合率は98%となった。検証に用いたデータには自転車を漕ぐ音や人の大声等の雑音が多く含まれていたにも関わらず、高い検出率と適合率が得られたので、本提案手法による拍手音の検出は、実用的なレベルに達していると判断できる。

拍手音の音高の分類手法の精度検証については、5人の被験者に3種類の指定した叩き方で各々100回叩いた時の検出率を調査した。その結果、DPマッチングによる識別手法は、中間の高さを含む3分類の識別タスクでは低い音の検出率が高く、高/低音の2分類の識別タスクでは、低い音の検出率が平均して75%以上となった。一方、高い音は被験者によって20%~90%の検出率を得た。以上の3分類のタスクで中間の音と高い音の検出率が低い原因は、図12に示す様な低い音と高い音の差が見られない状態に起因する。つまり、特定の被験者群は音高を分類できる様な拍手音を出せなかった結果と考えられる。

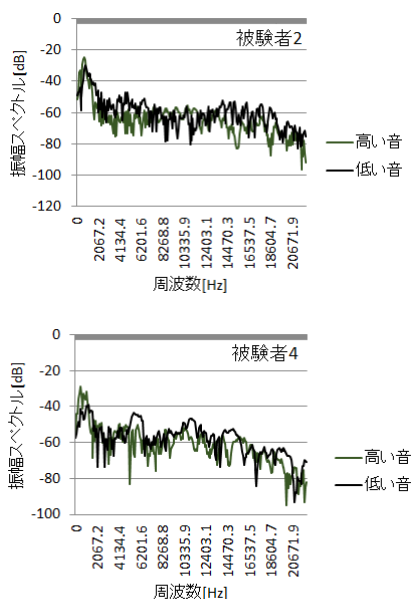


図 1 2. 音高調整できなかった2名の拍手音スペクトル

一方、音高調整ができたと考えられる被験者のデータを用いて検証した、音高の分類性能を以下の図13, 14に示す。

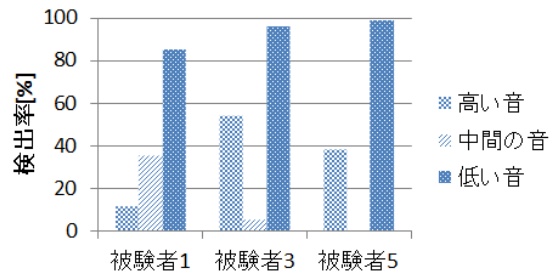


図 1 3. 3人の被験者に対する3音高の分類結果

図13に示す3音高の分類結果では、低い音の検出率が80%以上、特に被験者5は100%に近い検出率を得ているが、高い音と中間の音に関しては検出率が低いことが確認できる。この結果より、手の叩き方を指定しても高い音と中間の音は意図的な調整が難しいと考えられる。

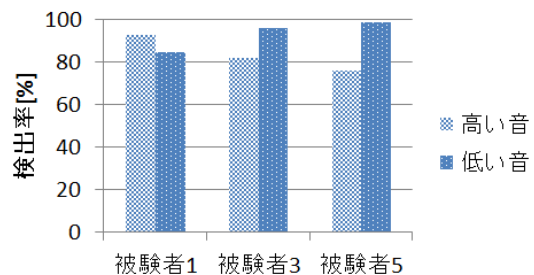


図 1 4. 3人の被験者に対する2音高の分類結果

図14に示す2音高の分類結果では、高い音と低い音の双方で70%以上の検出率が得られている。これは、中間音を分類対象から除外したことで高い音の検出率が増加したためであり、3分類では高音と中間音の混在が考えられる。以上の結果より、拍手音のコントロールができる被験者に対しては、提案した分類手法によって2種類の拍手を識別して操作に用いることができると考えられる。しかしながら、人によっては手を叩く音の音高を意図的に調整することの難しさも同時に明らかとなった。

今後の課題として、誰にでも手の叩き方によって意図的な操作が容易な音高以外の特徴量の抽出方法が考えられる。また、一般人にイルミネーションを意図的に制御させるために、手の叩き方の練習をさせるチュートリアル風な演出方法も検討に値するであろう。

## 5. 豊橋駅前広場での提示

「とよはしキラキラ☆イルミネーション」において、昨年(2015)の11月15日から本年(2016)の1月31日までの期間、豊橋駅前広場において本研究で開発したイルミネーション装置を展示した(図15)。

クリスマス前までの期間（12月13日から23日まで）と、それ以後正月までの期間（12月24日から元旦まで）に対して、拍手音を検知して再生された絵柄の数の1時間当たりの平均値を表1に示す。ただし、ツリーが点灯された午後5時から深夜0時までの時間帯での集計結果である。

表1. 検出された拍手音の数（回/時間）

計測期間	1人用ツリー	2人用ツリー
12月13日～23日	1 1 3	6 0
12月24日～元旦	1 0 9	6 6

拍手音によって再生される絵柄やパターンは、各々5秒間表示され、表示中は新たな拍手音を検出しない様に排他制御が実行される。したがって、本システムで検出可能な拍手の理論的な最大数は  $60 \times 60 / 5 = 720$  回/時間となる。この最大数と比較すると、拍手による対話操作が間断なく行われている状態ではないが、1人の拍手音だけに反応するツリー（表中の1人用ツリー）においては1分間に1.8回程程度の拍手音が検出されている。

ここで、拍手はしたものの音が小さかった場合や他の拍手音と重複した場合に検出されなかった拍手を考慮すると、対話的な操作を試みた人は表の値よりも多かったものと推察される。また、2人以上が同時の拍手した音に反応するツリー（表中の2人用ツリー）は、1人だけの拍手音には反応せずにその回数が記録されないのので、1分間に1回程程度という相対的に低い値となっている。実際に、駅前広場を通行する際にイルミネーションを鑑賞する総人数を考慮すると、多くの割合の人達が対話的な操作に興味を示し、拍手を試みたと考えられる。



図15. 豊橋駅前広場での展示風景[5]

## 5. おわりに

今後は、絵柄や文字を光源パターンに変換する機構の性能向上を目指す。特に、少数のLED光源から生成される点灯パターンで図柄や文字を正確に表示する事は、情報量の観点からも本質的には困難なので、例えば文字を表示する場合にはその視認性を最大化できる写像手法が必要となる。

現時点ではグラフィカル・モデル等を導入し、最適な点灯パターンを確率伝播モデルとして最適化する手法を検討中である。

セル・オートマトンやパーティクル・システムに基づく点灯パターンの自動生成は現時点ではハードコーディングされているので、パターン生成の共通部分を抽出し、振る舞いの拡張やカスタマイズが容易になる機構を開発する必要がある。

今後は、拍手に代わる手軽で直観的な対話的操作方法や演出パターンの設計支援方法も、娯楽性向上の観点から検討していく予定である。

## 謝 辞

本報告で紹介したイルミネーション装置は、平成27年度の豊橋市イノベーション創出等支援事業から開発費の支援を頂きました。開発に際して貴重な意見を賜りました、共同研究事業者の株式会社フカミの深見様、松村様、林様、および株式会社サイエンス・クリエイトの伊村様に、心より感謝申し上げます。

## 参考文献

- [1] 吉久光一, " 屋外の音の伝搬における空気吸収の計算, " 騒音制御, Vol.21, No3, pp130-135, 1997.
- [2] 相馬稔, 小川毅彦, 金田一, 森きよみ, " 短時間フーリエ変換による衝撃波系のスペクトル解析, " 電子情報通信学会総合大会, A-10-13, 2000.
- [3] 内田誠一, "[特別講演] DP マッチング概説 ～基本と様々な拡張～, " 電子情報通信学会, PRMU2006-166, 2006.
- [4] 伊藤智則, 上原聡介, 笠原弘樹, 市野将嗣, 吉井英樹, 鶴丸和宏, 甲藤二郎, 小松尚久, " 加速度センサを用いた歩行状態識別におけるケプストラム分析の有効性に関する一考察, " 電子情報通信学会, ICM2011-38, LOI2011-63, Jan, 2012.
- [5] 『朝日新聞デジタル』2015年12月8日「愛知）イルミネーション、拍手で七変化 豊橋駅東口」