

集団を対象とした マス・コンピュータ・インタラクション(MCI)の提案

Mass Computer Interaction (MCI)

上野 塁† 外村 佳伸‡
Rui Ueno Yoshinobu Tonomura

1. はじめに

現在, IT 技術の発展により, PC やスマートホンを始めとした様々な情報システムが普及している. これらの情報システムは主に使用者個人とのインタラクションを想定している. それに加え, 近年では大型ディスプレイを用いたデジタルサイネージのように, 公衆空間における情報システムの普及も始まりつつあり, 行き交う複数の人を対象とした情報提示に用いられ, 効果の検証が行なわれている[1].

今後, 情報システムの広スペース化, 公衆化が進むことで, 我々の身の回りの至るところに情報システムが配置され, 人々はそれらに対し, 当たり前のようにインタラクションをするようになると思われる.

我々はこのような身の回りの広い空間, いわば環境の一部となった情報システムのことを環境型情報システムと呼び, それと人との関係のあり方について研究している. 特にシステムと複数の人を対象に, どのようなインタラクションが可能か, またはあり得るかについて検討している.

本論文では, まず集団と環境型情報システムの特徴や, それらを用いた際の課題を挙げ, 集団とシステムとのインタラクション体系である Mass Computer Interaction (MCI) を提案する. また, プロトタイプシステムを構築し, 実験を行なった結果についても述べる. MCI を利用することにより, 人々のいる場を盛り上げるといったアミューズメントとしての活用や, 人々が起こす特定の状態をキーとしたセキュリティとして, また人々を特定の並び方になるよう促すシステムなどといった様々な応用が考えられる.

2. 集団とシステム

2.1 集団

集団とは, 人の集まりのことを指す. 集団を構成する個人に対し, [2]では公衆空間におけるディスプレイとインタラクションを起こす際の人側からのインタラクション・モダリティを定義している (図 1).

また, これらのインタラクション・モダリティは, それが自ずと醸し出してしまう暗黙的なものであるか, あるいは自らの明確な意思から表れる明示的なものであるかといった性質を持つ. モダリティの間に明確な境界はないが, 図 1 において上にあるものほど暗黙的, 下に進むにつれ明示的な意味合いが強くなっていく.

例えば, 歩行者の移動がシステムによって捕捉され, ライトが点灯するなどシステムによる何らかのアクションが起きた状況では, 人は自身の「位置」という暗黙的なモダリティを使用したことになる. 一方, 歩行者がシ

ステムに対し, 「ジェスチャ」や「タッチ」といった意図的なアクションをシステムに対して起こす場合は, 明示的なモダリティを使用していると言える.

implicit 暗黙的	Presence	存在
	Body position	位置
	Body posture	姿勢
	Facial Expression	表情
	Gaze	向き
	Speech	声
	Gestures	ジェスチャ
	Remote Control	リモート操作
	Keys	キー入力
	Touch	タッチ
explicit 明示的		

図 1: インタラクション・モダリティ

このような個人によるインタラクション・モダリティは, 様々な場所や状況において形成される集団に対しても考えることができる. 例えば, スタジアムの客席にいる人々は明示的な度合いが強い集団であり, 駅やスクランブル交差点にいる人々は暗黙的な度合いが強い集団である (図 2).



図 2: 明示的集団・暗黙的集団

この 2 つの集団の違いは, 集団を構成する個人ごとの関係性の深さに起因する. スタジアムにいる人々はある程度共通した目的を持っているのに対し, 駅を利用している人々はそれぞれ異なる目的を持つ. 集団に対して何かの指示をする場合, 明示的な度合いが強い集団に対しては比較的複雑な指示が可能である一方, 暗黙的な度合いが強い集団に対してはそれが難しいことが考えられる. そのため, システムを構築する場合は対象とする集団がどのような性質を持っているか, またどのような表現が可能かを十分に検討しておく必要がある.

† 龍谷大学大学院, Graduate school, Ryukoku University.

‡ 龍谷大学, Ryukoku University

2.2 集団を対象としたシステム

大型ディスプレイの普及やセンシング技術の発達により、広い場所にいる集団を意識した情報システムの研究が進められている。[3]では、群集の行動の分析・可視化を目的としたアルゴリズムの研究がされている。具体的には、群集の様子を撮影したビデオ映像から人の移動を追跡し、流体的に表現し分析をする。また、[4]のように集団パフォーマンスの支援を目的とした情報システムの開発も行なわれている。これは、人々に様々な色を発光させることのできる傘を持たせることで、全体として美しいアートの表現をするものである。

これらの例は、個別の目的や条件に特化した手法を用いてシステムの開発や提案がなされている。しかし、今後情報システムの環境化が進み、集団を対象とした情報システムの開発が増えることを想定すると、システムと集団のインタラクションの基本的な要素、処理の枠組みが必要である。そのために、集団によってどのようなアクションが表現可能か、またシステムが集団のどのような情報をセンシング可能なのかを検討することが、分野全体の研究加速化につながると考えている。

そこで本研究では、集団とシステムにおける情報のやり取りを体系的に考え、操作の基本形の定義として、Mass Computer Interaction (MCI) を提案する。

3. 集団とシステムのインタラクション

3.1 Mass Computer Interaction (MCI) の提案

従来、人とコンピュータとのやりとりを HCI (Human Computer Interaction) として、主に個々の人とコンピュータの関係を中心に研究開発が進められてきた。ここでは Human を Mass に置き換え、集団とシステムとのやりとりとして Mass Computer Interaction (MCI) を提案する。MCI における情報の流れを図 3 に示す。

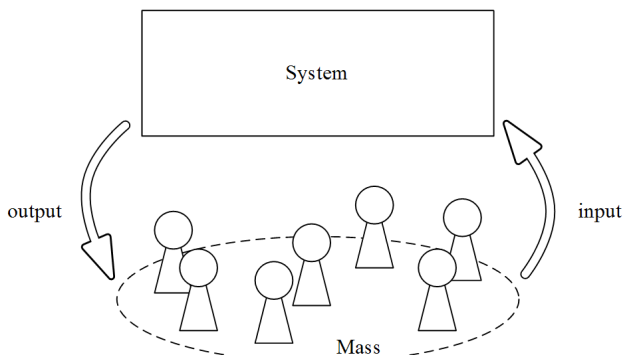


図 3: MCI における情報の流れ

集団は人の集まりである。一人毎のアクションの集まりが集団としてのアクションとなり、システムに入力される。入力を受けたシステムは、それを基にした処理を行ない、対象である集団に情報を出力する。この MCI モデルのうち、今回は集団のシステムに対する入力について主に検討する。

[2]では、個々の人がパブリックディスプレイに対して入力可能なインタラクション・モダリティを定義している。一方、集団によるインタラクション・モダリティを考える上では、まずは「タッチ」や「ジェスチャ」とい

った明示的なものよりも、暗黙的傾向の強いもの、すなわち人から自然と醸し出るものを使用する方が、様々な集団の状況に対応できると考えられるため、より暗黙的な「存在」および「位置」を主にインタラクションのためのアクションとすることとした。

集団のインタラクション・モダリティにおける「存在」とは、人々がいるか、いないか、ということである。それにより、空間における人の量、すなわち集団の密度という情報として解釈することができる。また、「位置」を用いることにより、人々の配置から形をつくるのが可能であり、また人々の位置の偏りから重心を求めることができる。「存在」「位置」から表現することのできる集団としての情報の例として、形、密度、重心に関する表現を図 4 に示す。

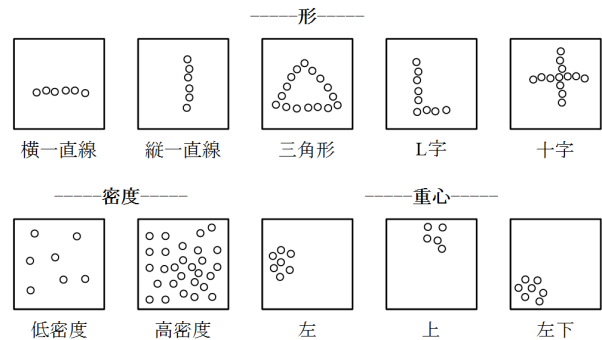


図 4: 「存在」「位置」による表現

一般に、空間の大きさや人々の行動可能範囲の広さによって構成される集団の規模は変化する。例えば、廊下という空間では数人オーダーの集団しか構成できないのに対し、コンサート会場では数百人以上のオーダーが構成可能である。集団とシステムのインタラクションを体系化するためには、様々なオーダーの人数に対して柔軟である必要がある。そこで少人数の集団をも想定するとともにシンプルな判定パターンとするために、上部から見下ろした空間を縦 3 横 3 のマトリクスに分割、その中で密度や重心を定義し、存在や位置と判定することとした。表現可能な情報は減少するものの、特徴的な形、例えば一直線や十字、X 字といったものであれば表現可能であること、密度に関しては 10 段階程度、重心に関しても後述するが 25 通り程度表すことができるため、このような手法を採用した。

以上を踏まえ、集団とシステムのインタラクションの基本形となるような Mass Computer Interaction (MCI) における入力としての解釈例を図 5 に示す。

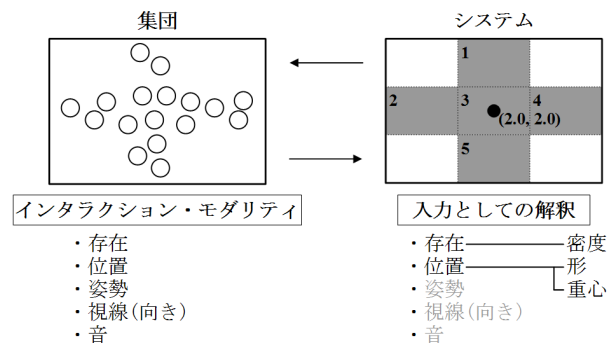


図 5: システムによる入力としての解釈

図5は、集団の存在と位置から密度、形、重心という3つの入力をシステムに対して行っている様子を示している。システムは、得られた入力から定義した3×3のそれぞれのマスをも0 (off) 又は1 (on) の状態にすることで、入力として解釈する。システムが人の存在を検出したと判断したマスは1 (on) となり、図中では色をつけて表現している。

3.1.1 密度

密度は0 (最低密度) ~9 (最高密度) の10段階で表現可能である。空間上に多くの人で埋め尽くした場合、高密度となり、人が少なければ低密度となる。図5の場合、1 (on) のグリッドが5マスであるため、密度は5となる。

密度は0~10の数値として表すことができるため、システムに対するパラメタとして活用可能である。

3.1.3 重心

様々な0 (off) 又は1 (on) の組み合わせにおいて、X軸およびY軸それぞれの1 (on) である座標の平均を求めることで、その状況の重心と定めた (図6)。具体的には、重心の座標を取り得る基準点を、それぞれのマス9点に加え、それらの間を補間する16点の合計25通りとした。この25点に当てはまらない座標を重心とした場合は、最も近い基準点にシフトする。左上の座標を(1.0, 1.0)とし、X軸Y軸それぞれの刻みは0.5である。

重心は座標として扱われるため、集団全体の持つ方向性として活用可能である。

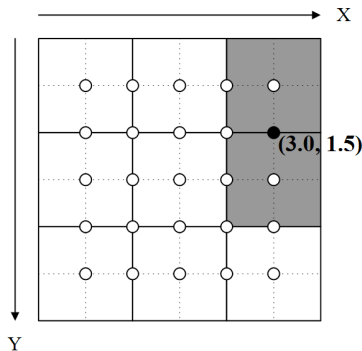


図6: 3×3マトリクス上で表現可能な集団の重心

3.1.2 形

0 (off) 又は1 (on) の状態を持つマスが9マスあるため、形の組み合わせの総数は $2^9 = 512$ 通りとなる。その中で、形状に比較的特徴を有している組み合わせを挙げたものを図7に示す。

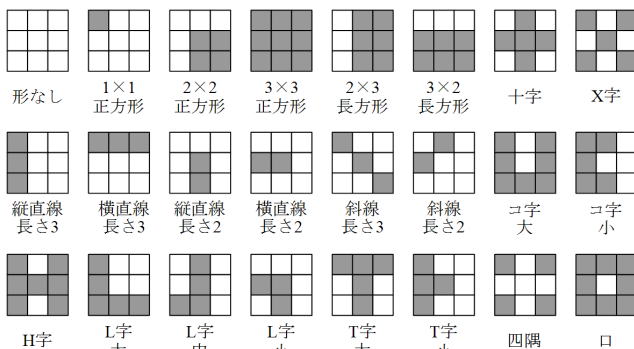


図7: 3×3マトリクス上で表現可能な集団の形

この他にも、S字や定義されていない形が考えられるが、複数人で形をつくるという行為を想定したとき、人が簡単に意識できる以上に複雑な組み合わせを省いて考えた。なお、形は特定の意味を持つキーとして活用可能である。

3.2 プロトタイプシステム

以上のMCIを踏まえ、プロトタイプシステムを構築した。具体的には、俯瞰設置したカメラにより集団のインタラクション・モダリティをセンシングし、目の前の大型ディスプレイに情報を表示させるシステムである。

3.2.1 システム構成

使用したハードウェアは、PC、Webカメラ (Logicool Webcam C930e)、複数の60"大型液晶ディスプレイ (SHARP PN-L600B) である。

Webカメラは天井の制約から地上から3.16mの高さに、空間を真下に見下ろすよう設置した。そのとき、カメラで撮影可能な横幅は5.28m (レンズ画角79.8°)、奥行は2.88m (レンズ画角49.0°)であった。図8に今回構築したシステムの外観を示す。



図8: プロトタイプシステム外観

3.2.2 集団情報センシング処理

本プロトタイプシステムのコントロールパネルを図9に示す。

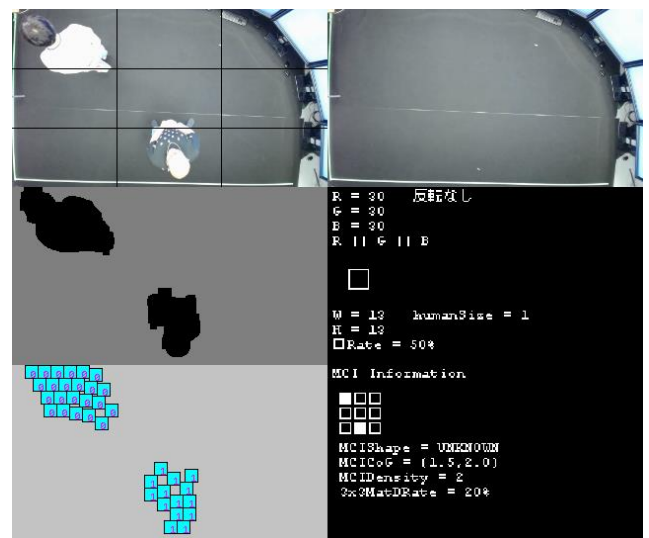


図9: コントロールパネル

最初に、空間上にいる人の検出を行なうための準備として、取得した画像から背景差分による2値化処理を行な

い、前景画像を取得する。図 9 左上はカメラが現在取得している画像に 3×3 グリッド線をオーバーレイしたものであり、右上はあらかじめキャプチャ済みの背景画像である。現画像と背景画像とで対応する画素の RGB 値のそれぞれの差を求め、一定以上の差が見られれば、その画素を図 9 左中のように黒く表示させる。パラメタの詳細な検討が必要ではあるが、今回は RGB いずれかの値が原画像と背景画像とで 30 以上の差があれば差分として検出することとした。

次に、取得した前景画像から人をモデルとして検出するためのブロック化処理を行なう。これはカメラの位置や高さによって条件が変わることを想定し、仮想的な存在検出単位としても受けた概念であり、ブロックのサイズは適宜変更可能である。カメラの設置位置によっては、人ひとりが空間の大部分を占める可能性があるが、その場合複数のブロックを一人に割り当て、疑似的に複数人として解釈し処理することとした。ブロック化処理は、設定されている大きさの矩形に画像上を走査させ、矩形内にある前景画素の量が閾値以上あれば、そこにブロックを配置する、というものである。図 9 の場合、ブロックは W=13, H=13 として処理を行ない、閾値は 50%とした。

最後に、3×3 で分割されたそれぞれのマトリクス上に存在するブロックの数を基に 0 (off) 又は 1 (on) を決定し、密度、形、重心を取得する。図 9 においては、各マトリクスにブロックが 20%以上存在する場合、1 (on) としており、図右下には現時点の集団の情報として以下の情報が表示された。

形 : UNKNOWN (未定義の形)
 重心 : (1.5, 2.0)
 密度 : 2

結果のうち形と密度を大型ディスプレイに表示させたものを図 10 に示す。長さ 3 の縦直線で密度は 3、重心は表示されていないが(3.0, 2.0)である。



図 10: 大型ディスプレイによるフィードバック

3.2.3 実験

以上のような手法で集団の情報をセンシングする際の検討課題として、カメラ画角に伴う被写体の映り方の画面内での違いについての問題がある。特に今回の実験環境のように広角のカメラで地上に近い場合に顕著である。そこで人の立つ位置によるブロック検出数について調べた。身長 170cm の男性を対象に、各マトリクス内で最も

大きく写る場所に立たせ、ブロックの数を測定した。図 9 を例に取ると、画面端に立っている者は身体が斜めに引き伸ばされたように写るため、マトリクス毎で異なる閾値を設定しなければならないためである。なお、最も大きく映る場所は、そのマトリクスの隅に頭部が入る限界の位置であり、各四隅に関しては 1 箇所、上下左右端は 2 箇所、中央は 4 箇所考えられる。それらの値の平均値と、中央を 1 とした際の中央との比率を図 11 に示す。

Display		
38 (2)	32 (1.68)	38 (2)
32 (1.68)	19 (1)	32 (1.68)
38 (2)	32 (1.68)	38 (2)

図 11: マトリクス毎の最大ブロック平均値と比率

次に、各マトリクスの 0 (off) 又は 1 (on) を決定する適切な閾値について検討するための実験を行った。

実際に集団に特定の形を形成してもらい、マトリクス検出閾値を 0% (エリア内に少しでもブロックが存在) から 100%までの 10%刻みで変更させ、システムが正しい形を検出できているかという方法で実験を行った。集団は 20 歳代の男性 3 人、女性 2 人の計 5 人であり、図 12 のそれぞれの形をとるよう指示をした。これらの形は、on のマトリクスが最大 5 つまでで形成可能なものである。一人では複数のマトリクスを on の状態にすることはできないと仮定したため、これらの形を対象とした。3×3 で表現できる形のうち、on 状態のマトリクスがおよそ 5 つあれば大部分の形状は表現可能であると考えたためである。また、実験中はディスプレイによる情報の表示はしていない。地面にはカメラの撮影可能な範囲のみを示すためラインテープを使用し、マトリクス毎のグリッドは示さないこととした。

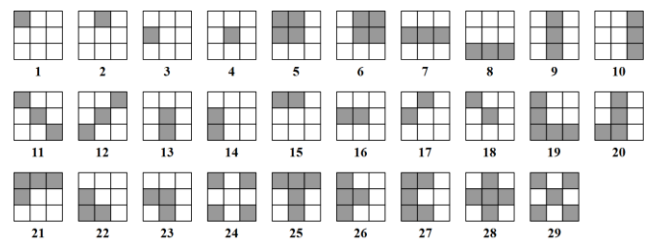


図 12: 集団に指定した形

実際に図 12 における 7 番 (横直線 長さ 3) の形を指示した際の集団の様子とその検出結果 (閾値 20%) を図 13 に示す。

そして、形とマトリクス検出閾値の関係を表したものを表 1 に示す。横軸が形のインデックス番号、縦軸が閾値であり、0 が検出失敗、1 が成功である。

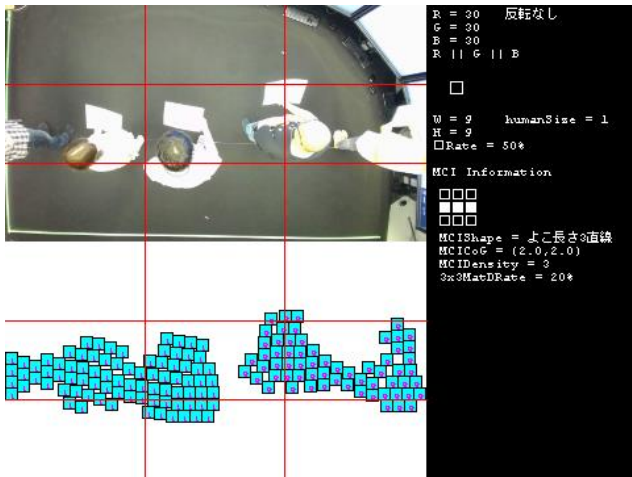


図 13: 実験の様子

表 1: 形と検出閾値の関係 (重み係数なし)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0%	0	0	0	0	0	0	0	0	1	0	0	0	0	0
10%	0	0	0	0	0	0	0	1	1	1	0	0	0	0
20%	0	0	0	0	1	1	1	1	1	1	0	0	1	1
30%	1	1	0	0	0	0	1	0	1	1	0	0	1	1
40%	1	1	0	0	0	0	1	0	1	1	0	1	1	1
50%	1	1	1	0	0	0	1	0	0	1	0	0	1	1
60%	1	1	1	1	0	0	0	0	0	1	0	0	1	1
70%	1	1	1	1	0	0	0	0	0	0	0	0	1	0
80%	1	1	1	1	0	0	0	0	0	0	0	0	0	0
90%	1	1	0	0	0	0	0	0	0	0	0	0	0	0
100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0

結果として、マトリクス検出閾値が 20%のとき、最も検出成功率が高く、計算すると 62%となったものの、形によっては検出できていないものも多い。これは、比較的マトリクス検出閾値の低いときに顕著に見られる。少しでもマトリクスをブロックがまたいでしまうと誤検出が多く発生していると思われる。

そこで、図 11 の結果から得られたマトリクス毎のブロックサイズの比率を基に、四隅と上下左右端のマトリクス検出閾値にそれぞれ重み係数をつけることで、ブロックの越境による誤検出の低減を試みた。検出率が 20%の場合、中央はそのままの 20%で判定を行ない、四隅は重み係数 2 を乗算した 40%、上下左右端は 1.68 を乗算した 33.6%となる。重み係数を用いた検出結果を重ね合わせ、改善箇所をハイライトしたものを表 2 に示す。従来の検出結果において、越境による誤検出が発生していた箇所が、重み係数を用いた検出を重ねあわせることで比較的改善されていることがわかる。

これらの結果から、マトリクス検出閾値毎の検出成功率を表したものを図 14 に示す。結果として、従来の検出方法では閾値が 20%のとき、62%の検出成功率であったが、

重み係数による検出を重ねあわせることで、検出成功率が 72%と上昇した。ただこの値は安定した検出のためにはまだ十分と言えない。特に、斜線、コ字、十字、X 字の形については検出率がかなり悪く、このままの条件では検出できないことがわかる。

表 2: 形と検出閾値の関係 (重み係数あり)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0%	0	0	0	0	0	0	0	1	0	0	0	0	0	0
10%	0	0	0	0	1	1	1	1	1	1	0	0	0	0
20%	1	1	0	0	1	1	1	1	1	1	0	0	1	1
30%	1	1	1	0	0	0	1	0	1	1	0	0	1	1
40%	1	1	1	1	0	0	1	0	1	1	0	1	1	1
50%	1	1	1	1	0	0	1	0	0	1	0	0	1	1
60%	1	1	1	1	0	0	0	0	0	1	0	0	1	1
70%	1	1	1	1	0	0	0	0	0	0	0	0	0	1
80%	1	1	1	1	0	0	0	0	0	0	0	0	0	0
90%	1	1	0	0	0	0	0	0	0	0	0	0	0	0
100%	0	0	0	0	0	0	0	0	0	0	0	0	0	0

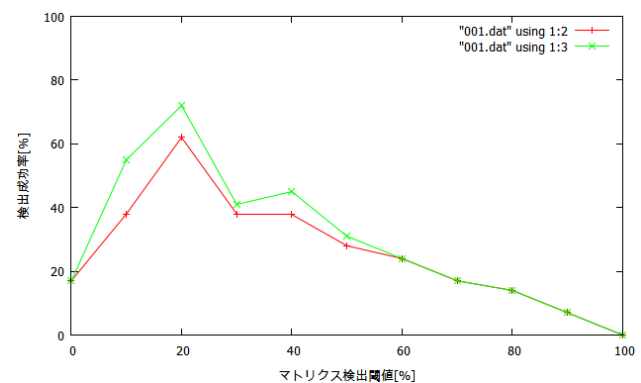


図 14: マトリクス検出閾値毎の検出成功率

対策として、カメラをより高い位置に設置し、一人あたりの影響力を小さくすることや、レンズキャリブレーションとして分割境界を最適化するなどによる誤検出の低減が考えられる。

3.3 応用例 PassMass

提案した MCI を基に、ひとつの応用例として PassMass というアプリケーションを作成した。これは、緩やかなセキュリティとしての活用例である。図 15 に PassMass を利用している様子を示す。

これは、システムに対し MCI における入力を利用するもので、図 15 の例では密度を利用し、一定の密度を保つことでロックを解除するようなアプリケーションである。ディスプレイにはゲージと正解の密度値の範囲を表示させ、現在の密度と正解との関係を利用者に提示している。

なお、密度モードの他に形、重心モードも存在する。複数のモードを組み合わせ、ディスプレイ表示によるフィードバックを使用しない設定にすることで、より複雑で強固なセキュリティとしての活用も考えられる。



図 15: PassMass

4. おわりに

本提案では、ますます広スペース化する情報システムの普及に際し、集団とシステムを対象としたインタラクション体系 **Mass Computer Interaction (MCI)** を提案した。

しかし、カメラの高さ、人の位置によるカメラ写りの変化に対し完全な対応ができなかったため、所望のパターンを全て安定して取得できるところまでいかなかった。今後、重み係数の適切な決定や、カメラの歪みを考慮したマトリクス分割についての検討を行わなければならない。

また今後は、存在や位置以外の新たなインタラクション・モダリティについても検討していきたい。例えば環境音を取得することで集団の秩序度としての解釈が可能であると思われる。また、現段階では検討しきれていない、集団とシステムの関係性について考えたい。**PassMass** は、システムが集団に対して入力を促すシステム主導のものであると考えられるが、他にもパフォーマンスの補助としてシステムを利用する集団主導のものや、集団とシステムが協調的に物事を起こすようなものについても検討することで、新たな集団とシステムのインタラクションが生まれるのではないかと考えている。

参考文献

- 1) Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, Florian Alt: “Looking Glass: A Field Study on Noticing Interactivity of a Shop Window”, CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.297-306, (2012)
- 2) Jörg Müller, Florian Alt, Albrecht Schmidt, Daniel Michelis: “Requirements and Design Space for Interactive Public Displays”, MM '10 Proceedings of the international conference on Multimedia, pp.1285-1294, ACM New York(2010)
- 3) Brian E. Moore, Saad Ali, Ramin Mehran, Mubarak Shah: “Visual Crowd Surveillance Through a Hydrodynamics Lens”, Communications of the ACM, Vol. 54, No. 12, PP.21 64-73, (December 2011)
- 4) MIT Computer Science and Artificial Intelligence Laboratory. “MIT Members Invited to Color Nighttime Sky with Pilobolus & UP: The Umbrella Project”. May 8, 2013. <http://www.csail.mit.edu/node/1945>, (accessed July 25, 2014).