

UtteractiveWall : 声を集めてインタラクティブ化するアンビエントスペース

山本 直慶[†] 外村 佳伸[†]

本論文では、人の関心を引き起こし、インタラクシオンを誘発する環境的な情報空間（インタラクティブ・アンビエントスペース）として、人の声に着目した UtteractiveWall を提案する。UtteractiveWall は、壁型のタッチインタフェースを持つ複数の大型ディスプレイにより構成し、人の発する声を次々と録音し、それをオブジェクトとして可視化し、さらにそれに触れると録音した声が再生されることで、人々の関心とインタラクシオンを生み出す空間を作り出すものである。カメラとマイクを協調的に用いることで人々の位置を判定し、人々の動きに応じてオブジェクトがダイナミックな動きを示す特徴を持つ。本システムで構築した環境は、アンビエントスペースを用いた今後の様々なインタラクシオンの基本環境として展開が期待できる。インタラクティブ発表用にはデスクトップ型を展示する

UtteractiveWall : Ambient Space for Collecting, Showing, and Replaying People's Voice

NAOYOSHI YAMAMOTO[†] YOSHINOBU TONOMURA[†]

This paper proposes “UtteractiveWall” (Utter + attractive + interactive + Wall) which attracts people and stirs up them to join what happens there, what we call interactive, ambient space. UtteractiveWall catches the eyes of passers-by with bubbles floating on its large-and-wide display, when the person approaches the wall and touches one of the bubbles, a recorded someone's voice is replayed. Prompted person can record his/her voice which becomes one of the bubbles. This repetition enhances the ambient space to be more attractive. The system comprises seven units, each of which has a PC, a large vertical-type display with embedded touch sensors, a webcam, and a microphone. The proposed system could be a platform of future interactive applications on ambient space.

1. はじめに

我々の身の回りでは、パソコンに始まり、携帯電話、そして最近ではスマートフォンやタブレット端末が普及するなど、パーソナルな情報環境が充実しつつある。特にタッチインタフェースが本格化することで利用できるインターフェースの幅も広がってきたと言える。

一方で、大型ディスプレイの普及により、様々な公共の場所でネットワークに繋がれたディスプレイを用いたデジタルサイネージの活用が増えており、多人数向けに自在でダイナミックな情報や存在感のある映像が表示されている。しかし、人々がインタラクティブに関わるデジタルサイネージはまだまだ少ないと言える。

我々は、将来型の情報環境のひとつとして、公衆の目に触れる壁と人々とのインタラクシオンに焦点を当て、どのような情報環境が有望で、どのようなインタラクシオンがあり得るかを検討している。

本論文では、タッチパネル、カメラ、マイクなどのセンサーと、スピーカーを持つ複数の大型ディスプレイにより構成する環境を使用し、人の関心を引き起こし、インタラクシオンを誘発する情報空間（インタラクティブアンビエントスペース）のひとつとして UtteractiveWall の提案を行う。

UtteractiveWall は、人の発する声を次々と録音し、オブジェ

クトとして可視化し、さらに人のタッチにより録音した音声再生することを通して人々の関心とインタラクシオンを増強的に生み出し成長する空間を作り出すものである。

2. 関連研究

Prante ら公共空間における大型ディスプレイを用いたインタラクティブな空間に関しては、124個の LED で構成された 1.8×2.0m のディスプレイを用いて、利用者に気付きを与える Hello. Wall という試みを行っている[1]。この他にも、公共空間における人々の気付きとインタラクシオンについて様々な試みが行われている[2][3][4][5]。このように、大きなスケールで実現される状況は、複数の人々を巻き込む情報環境の“空間化”と言える。また、このような情報環境は、従来のパソコン利用などにおける目的志向型のパーソナルなインタラクシオンではなく、人々がシステムの表現に気付くことから始まる気づき誘発型インタラクシオンシステムという点で本論文と共通する問題意識を持っている。

また、平田らは大型ディスプレイを並べて人を囲み、遠隔にいる利用者同士があたかも同室にいるような感覚を伝えることの出来る空間として t-Room[6]を提案している。これは、コミュニケーション環境を意図して、人々の視野をおおう映像空間を構成したものである。本論文では、コミュニケーション環境ではなく、人々が関心を持ちシステムに近づいてきてから生まれるインタラクシオンの空間の構成を目的としている。

[†] 龍谷大学理工学研究科情報メディア学専攻
Graduate School of Science and Technology, Ryukoku University

3. UtteractiveWall

本論文で提案する UtteractiveWall とは Utter (声を発する) と attractive (人を引き付ける) さらに interactive (相互作用) の意味を含み、人の声を発する声 (つぶやき~叫び) を受け止め、それを可視化し、されにそれを人がタッチすると録音された声を再生することを基本とするシステムである。本章では、この UtteractiveWall のコンセプトと設計方針について述べる。

3.1. 関心とインタラクションの距離関係

人々が行き交う公共空間で、人々に関心を持たせインタラクションを誘発するような環境構成の方針として次のようなものを挙げた。

- ① 人に気付きを与える
- ② 気付きをきっかけにさらに人を引き込む
- ③ 人が思わずタッチするよう誘導する
- ④ 人が関心を持つ現象を起こす
- ⑤ 人に自らの参加を促す
- ⑥ 上記のポジティブスパイラルが起こる

Streitz らの Different Zone of Interaction[7]では1つのウォールの前に3重のコミュニケーションゾーン (人の位置に応じた処理ゾーン) をもうけている。本システムにおいても人の位置に応じて処理を適応させることを考えるが、我々の目的に応じた独自の景観距離、対面距離、タッチ距離の3つの距離関係に分類を行っている。

景観距離とは、人が景観として周囲を見渡す中で気づきを生む距離と考え、本システムではアンビエントウォール全体を見渡せる距離と定義した。対面距離とは、人が他の人や何らかの表示と対面し強く関心を抱く距離とし、本システムではディスプレイの表示内容をしっかり把握できる

距離と定義した。タッチ距離とは、人がディスプレイにタッチしやすい距離と定義した。各距離と視野の範囲の概念を図1に示す。実線は人から見たときの各距離の視野の範囲、点線はシステムの一つのカメラから見た各距離にいる人を捉えることの出来る視野を示す。

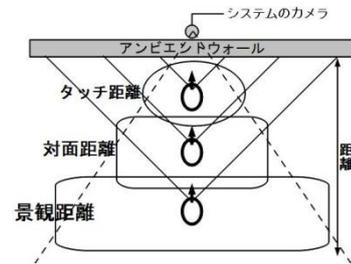


図1：距離関係

3.2. 人と UtteractiveWall 間のインタラクション

UtteractiveWall を実現するために、人の行動に対して UtteractiveWall が行う処理の流れを上記の距離関係において示したのが図2である。

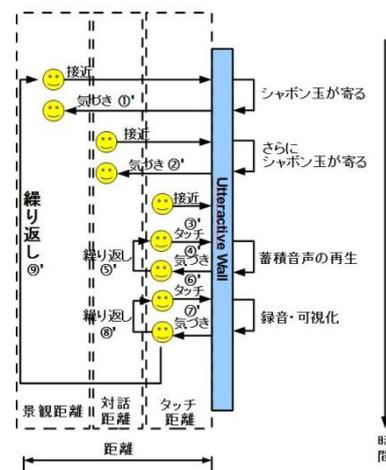


図2：声によるインタラクションの流れ

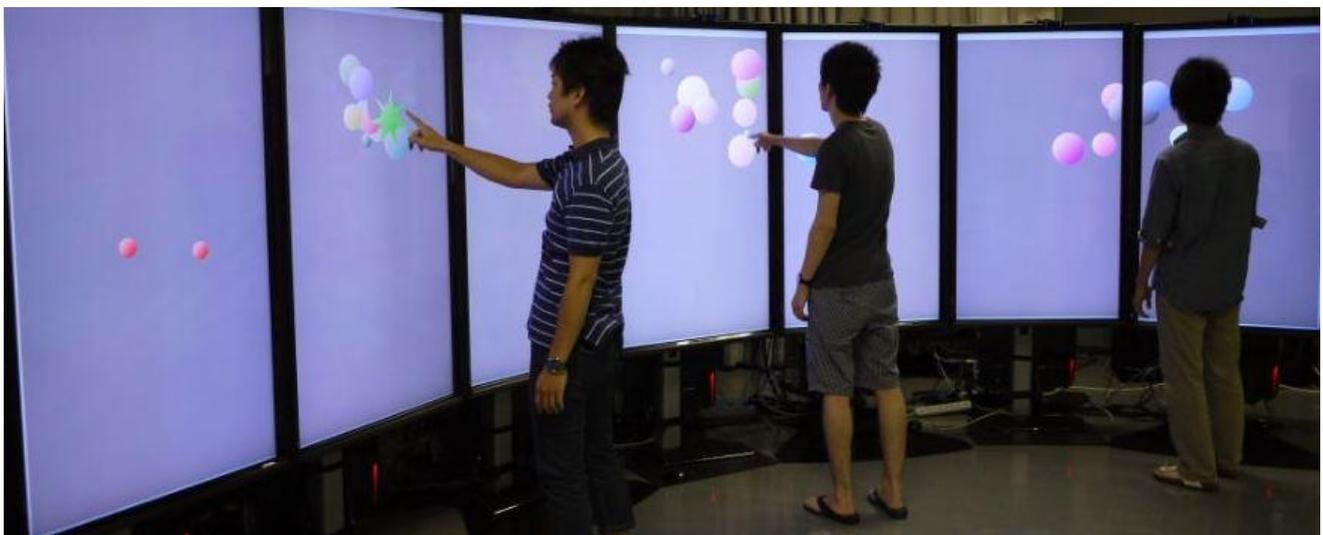


図3：UtteractiveWall の全体像

まず、人が景観距離まで接近すると、多くの円（本研究ではシャボン玉としている）がふわふわ動いていることに気付く（①）。各シャボン玉は人の声などが録音され可視化されたものである。これに人が興味を持ち、景観距離から対面距離まで近づく。そして、シャボン玉がさらに人に近づいてくる事に人が気づく（②）。さらに、タッチ距離まで人が近づきシャボン玉にタッチすると声が再生される（③）（④）。具体的には、この声を聞いた人が触発されて何か話して録音すると、また新しいシャボン玉が生まれる（⑤）。録音は、シャボン玉がない場所をタッチすることで開始され、人が話すときの声の大きさに対応したシャボン玉が生成される（⑥）（⑦）。以上のサイクルが続くと画面にシャボン玉がたくさん浮かぶようになる（⑧）。このシャボン玉が多人数に気づきを与え、新たな関心を呼ぶこととなり、掲示板のように成長する蓄積が生まれる（⑨）。

UtteractiveWall を実現するために技術的に重要なことは、人とシステムとの距離関係をつかむこと、すなわち人の存在と位置を把握することである。これにより、それぞれの距離におけるインタラクションのモードを変化させることができる。人の位置判定は、特にシャボン玉を人に近寄せるために必要となる。

4. プロトタイプシステム

以上の方針に基づき、UtteractiveWall のプロトタイプシステムの設計・実装を行った。システムが動作している全体像を図3に示す。複数の人がそれぞれにシステムに触れている様子を表われている。

4.1. 基本ユニットの構成

UtteractiveWall の基本プラットフォームは、タッチインタフェースを搭載した大型液晶ディスプレイにマイクと Web カメラをディスプレイ上部に、スピーカーと PC を下部にセットしたユニットとし、これを複数並べることが出来るものである。図4は、ユニットの寸法と設置状況を表している。

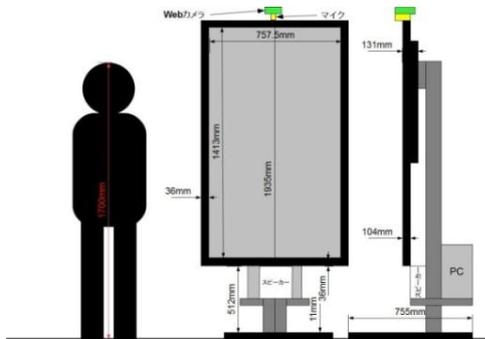


図4：ユニット構成図

具体的には 60 インチのタッチディスプレイを、スタンドを用いてディスプレイの下端が人の膝上の高さで、上端が人の頭上を超える高さ(1935mm)になるように設置しており、ほぼ等身大全身を表示可能としている。このユニットを複数並べることにより、複数のタッチパネル、カメラ、マイクによる入力と、横広がりの大画面表示が出来るディスプレイ、スピーカーによる音出力を有する空間的な環境を作り出すことが出来る。

4.2. システム構成

システムは、上記のユニットを7台隣接して設置し、全体で横広がり大きな壁になるように構成している。さらに、ホスト PC は、LAN 経由で全体を制御する。この全体の情報を共有できるようにネットワーク HDD を LAN 上に持つ。また、各ユニットのディスプレイにはユニット毎の出力は通常は表示せず、ホスト PC のディスプレイ出力による表示を行う。また各ユニットのタッチパネルによる入力情報はホスト PC で集約する。図5に各ユニットの物理的な配置図とネットワーク構成を示す。

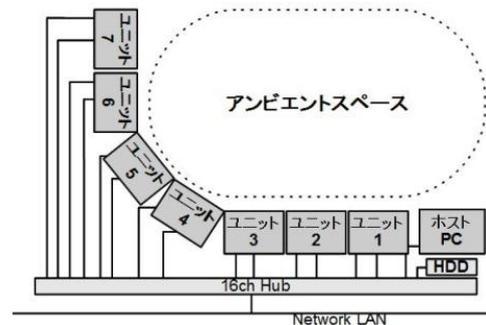


図5：各ユニットの配置とネットワーク構成図

各ユニットの配置は、公共空間の様々な場面を想定し、平面的に設置したユニット 1, 2, 3 と、ほぼ 90 度になるようにユニット 6, 7 を設置し、それらの間を滑らかなカーブになるようにユニット 4, 5 を設置した。ユニット 1, 2, 3 は横並びで平面に設置することで、横長の大型ディスプレイとして機能する。ユニット 3, 4, 5 では三面鏡のように設置することで、ユニット 4 の正面に立つことで左右から囲まれるような空間をつくっている。そして、ユニット 6, 7 と合わせてユニット全体では、公共空間の通路のコーナー部分を演出しており、様々なタイプの公共空間を圧縮したような環境となっている。

4.3. ソフトウェア構成

UtteractiveWall のソフトウェア構成は、各ユニットとホストユニットで分かれている（図6）。

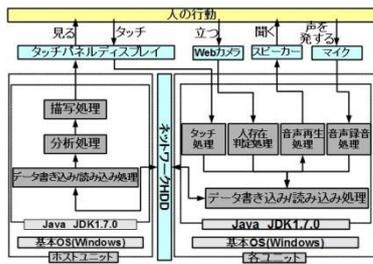


図6：システム構成図

各ユニットでは、Webカメラでとらえた映像から人存在判定処理を行い、判定した対応ユニットにおいてマイクからの音声を録音処理しネットワークHDDに書き込みを行う。また、再生指示が来た場合、再生処理を行い対応ユニットのスピーカーから音声を流す。

ホストユニットでは、各ユニットからの情報を集約し、そのデータを元に以下に示す描画を行う。

4.4. ユーザインタフェースの構成

UtteractiveWallのユーザインタフェースとして、録音した音声の存在を表すオブジェクト（シャボン玉）の例を図7に示す。

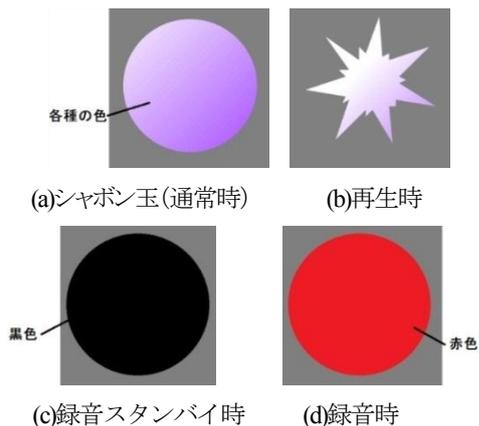


図7：UtteractiveWallのGUI

まず、録音音声を示すシャボン玉としては(a)に示すような円形（色つき）として表現する。(c)の黒色のシャボン玉は、シャボン玉がない場所をタッチした時に生成されるもので、録音スタンバイ状態を示す。次に、(d)のシャボン玉は、黒のシャボン玉を再度タッチすることで録音が始まり赤色に変化したものである。録音は一定時間（現在は5秒）行われ、録音が終了した時に(a)のシャボン玉が生成される。シャボン玉の大きさは録音時の声の長さに対応している。色は現在、人が見てある程度区別が付きやすいように12色に限定しその中からランダムに選んでいる。シャボン玉は壁の中でふわふわ浮かぶ。この生成されたシャボン玉を人がタッチすると(b)の状態に変化し、録音された音声が再生される。また、人の近くにあるシャボン玉が、その人の位置に集

まるようなアルゴリズムを組み込んでおり、人の位置に対しては、正面から左右にディスプレイの横半分の範囲までひきつけるようになっている。複数の人がいる場合、ひきつける力が大きい場所に行くようにしている。図8は人に引き付けられたシャボン玉の様子を示したものである。

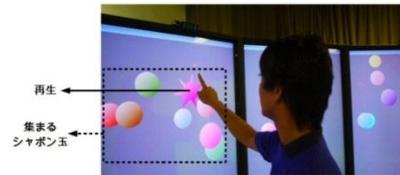


図8：近づくシャボン玉

生成されたシャボン玉は、人が近くにいる場合は、人の近くによって来るが、人が近くにいない場合は、シャボン玉が生成された位置（ホームポジション）に戻っていく。

4.5. 人存在判定処理

前述で定義した距離関係を判断するために、人存在判定システムを作成した。これを判定している状態を図9に示す。

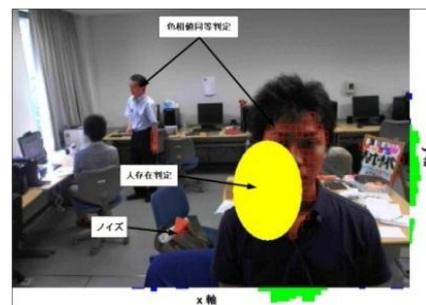


図9：人存在判定システム

本システムの目的から、人の存在に対して、人の特定や人数などの正確な状況はつかめなくても機能することから、処理としてはシンプルな色判定を用いることにした。また、人以外の対象も含めた今後の汎用性も考え、判定色は利用に当たって事前に決めることが出来るようにした。具体的には事前に、判定したい物体を画面上でタッチして指定し、その場所の示す色相を検知するものである。UtteractiveWallでは、まずデフォルト値として人の肌色を検知している。

画像処理では、Webカメラでとらえた映像を10×10ピクセルのブロックに区切り、その範囲での平均の色相値と取得した色相値に近い値の場合は、その範囲を抽出する（色相同等判定）。

人存在判定は大まかな位置とともに行う。これは、色相同等値判定を行ったブロック数をx軸、y軸で足しこんだ個数に対応した量を計算し（画面左部と下部に表示）、これが指定した閾値以上になった場所を抽出する（存在を黄色の楕円）。この時、色相同等判定で発生したノイズ部分は量が少ないので、人存在判定には影響が表れにくいと判断した。

5. システムパラメータの実験的設定

今回の提案手法において定義した距離関係の実値を実験によって求め、タッチ距離において適切に音声を録音できる設置条件を求めた。さらに存在距離を判定する手法を実験的に検証を行った。

5.1. Uteractive Wall と人の距離関係

距離関係（景観距離、対話距離、タッチ距離）を調べるために測定を行った。測定方法として景観距離については、ディスプレイに風景写真を全画面に表示し、風景写真の上端から下端まで首と目を動かさずに全体を見ることの出来る位置に被験者が立った距離とすることとした。対話距離は、タッチディスプレイに等身大の人の写真を表示し、被験者が写真の人物と対面するときに適切な間合いと思う距離とする。タッチ距離は、ディスプレイに人がタッチしやすい距離とすることとした。これを男 9 名、女 2 名の計 11 名を被験者として実験を行った。実験を行う前に、被験者にマイク角度を求めるために必要な身長と口から頭上までの距離の測定を行った。結果を図 10 に示す。

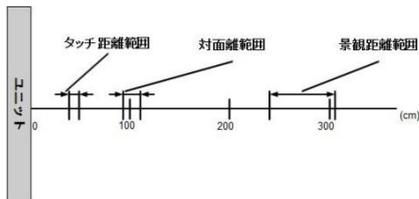


図 10: 距離関係図

景観平均距離273.7cm, 景観距離幅244.3~303.3cm. 対面平均距離 102.2cm, 対面距離幅 94.4~111.6cm. タッチ平均距離 47.1cm, タッチ距離幅42.9~51.2cmとなった。

以降の実験では、この各距離を用いて俯瞰カメラでの人判定を行う。

5.2. 俯瞰カメラによる人の距離判定手法

各ユニットの Web カメラをディスプレイの上部に設置したことで、撮影画面は俯瞰画像となる。この時、Uteractive Wall が正常に作動するための視野範囲として、日本の成人平均身長である男性170.7cm, 女性157.9cm で、タッチ距離～景観距離で顔が映る必要がある。この位置関係を図11に示す。

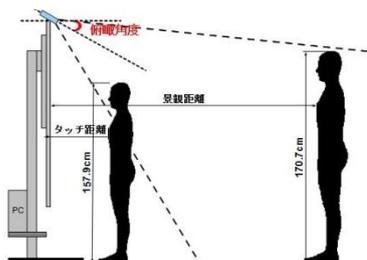


図 11: 画角測定図

これが成り立つために、平均身長である男女を被験者とし、タッチ距離で顔が映り、かつ景観距離で顔が映る画角を測定した結果、カメラの設置角度が水平から下向き15°の時、目的の画角を得られることが分かった。

この結果より、すべてのユニットの Web カメラを設置し人存在判定システムが各距離で人の存在を読み取ることができるか確認を行った。その結果を図12に示す。



(a)景観距離

(b)対話距離



(c)タッチ距離

図 12: 人存在判定

(a), (b), (c)は、それぞれ被験者が景観距離、対面距離、タッチ距離に立った時の判定結果である。この結果からそれぞれの距離で人の位置を読み取ることができた。また、景観距離では判定円が上部に表示されており直径が小さく、対面距離では、判定円が中部に表示され直径が景観距離の時よりも大きくなっていることがわかる。タッチ距離では、黄色い円が下部に表示され直径は最大になっている。この結果から、動作確認ができたとともに、俯瞰カメラであることにより画面中の上下位置と判定円の大きさで、人の位置が大まかではあるが簡単に把握できることが分かった。

6. 考察

前章での各種設定値、処理手法を用いて、プロトタイプシステムを動作させ動作確認を行うとともに、多くの人に体験してもらった。

6.1. 処理性能

画像処理や音声処理等は、各ユニットの PC で分散して行われ、ホスト PC では、それらの結果を集めて処理する構成のため、各画面に1人、すなわち全体で7人が同時にインタラクションをしてもリアルタイム処理としては全く問題なく動作した。ただし、音声データを含め制御信号のやりとり

を、LAN 上の共有ディスクを介して行っているために、わずかなディレイ（コンマ数秒の遅れ）が生じた。この点については、高速化の様々な工夫が可能なため今後取り組む。

6.2. ユーザインタフェース

シャボン玉の表現方法については、録音時・再生時の表現方法を工夫することで、シャボン玉の状態や動作状況を判断できるようにした。特に、録音時の声の長さに応じて、シャボン玉の大きさを変えたことで、たくさんのシャボン玉が表示されている際にメリハリが効いた画面になると同時に、録音メッセージに関する長さとしての表現となっているメリットがある。またシャボン玉の色については、今回ある程度の区別が付きやすい12色に限定したが、自分の録音したシャボン玉の区別が付きにくい、との利用者意見もあり、匿名的ではあるが自分にはわかる、個性的なシャボン玉表現を検討したい。

6.3. 人存在判定について

人存在判定システムでは、Web カメラで俯瞰することで本システムの目的に照らしてさらに機能するシンプルな判定が実現できた。その結果、シャボン玉を人に近寄せられるなどのダイナミックな動きも可能になった。その一方で、シャボン玉が人のいる一か所に寄りすぎて目的のシャボン玉にタッチしにくい状況が発生したため、今後改善していく。さらに、現在、事前に処理する色の指定と簡単なキャリブレーションが各ユニット毎に必要なため、今後その簡略化も必要である。

6.4. 利用者評価

大学における公開デモンストレーションとして多くの人に体験してもらったが、簡単に利用例を示すだけで喜々としてタッチし、録音する姿が常に見られた。

しかしながら、本来のコンセプトに基づく利用者評価には、気づき方、誘導性、アクションの誘発性などをシャボン玉の数や表現方法も絡めた評価が必要となり、現在その評価法の検討中である。

7. UtteractiveWall mini

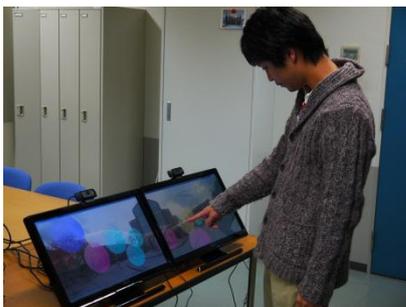


図 13 : UtteractiveWall mini 動作図

インタラクティブ発表では、本提案のオリジナルのシステムは設置できないため、デスクトップサイズのシステムを構築した。フル HD 表示の液晶ディスプレイを2~4面並べて、同様の処理を行うものである。距離関係や人存在判定処理上で基準は変わるものの、基本的なコンセプトは表現できている(図13)。

8. さいごに

本報告では、人とシステムとの距離関係に着目し、人の声を記録したシャボン玉をきっかけに人を引き寄せさせる空間を作り出す UtteractiveWall を提案した。また、システム構成上の様々な条件の設定値についても実験的に検証しながら実現した。

今後の展開としてシャボン玉の表現に関し、色や形に何かしらの情報表現を、直視性を損ねない範囲で取り入れることで、より表現力のある魅力的なものにできると考えている。さらに、人側の情報として、存在と位置だけでなく、意識的/無意識的に示す情報を検知することで、例えば人自身や目的にふさわしいシャボン玉を特定の人に引き寄せるなどが可能になる。それと合わせて、声以外の情報をシャボン玉に託すことへの展開も検討している。

参考文献

- [1] Prante, T. et al. Hello.Wall – Beyond Ambient Displays. Adjunct Proceedings of UbiComp 2003, pp277-278
- [2] Norbert A. Streitz, Carsten Röcker, Thorsten Prante, Daniel van Alphen, Richard Stenzel, Carsten Magerkurth: “Designing Smart Artifacts for Smart Environments” IEEE Computer Magazine (March 2005) pp.41-49
- [3] Patrick Tobias Fischer, Eva Homecker “Urban HCI: Spatial Aspects in the Design of Shared Encounters for Media Facades” CHI 2012, pp307-316
- [4] Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, Florian Alt: “Looking Glass: A Field Study on Noticing Interactivity of a Shop Window” CHI 2012 pp.297-306
- [5] Maurice ten Koppel, Gilles Bailly, Jörg Müller, Robert Walter: “Chained Displays: Configurations of Public Displays Can Be Used to Influence Actor-, Audience-, and Passer-By Behavior” CHI 2012 pp.317-326
- [6] Keiji Hirata et al., : t-Room: Next Generation Video Communication System. In Proceedings of World Telecommunications Congress 2008 (with IEEE Globecom 2008).
- [7] N. Streitz et al., “Ambient Displays and Mobile Devices for the Creation of Social Architectural Spaces: Supporting Informal Communication and Social Awareness in Organizations,” *Public and Situated Displays: Social and Interactional Aspects of Shared Display Technologies*, K. O'Hara et al., eds., Kluwer, 2003, pp. 387-409.