

人の位置移動による状況即応型 デジタルサイネージインタフェースの構成法

木原 民雄[†] 渡辺 浩志[†]

最近では、駅や商業施設などの公共空間に設置されるデジタルサイネージの数が増加している。場の状況を捉えるセンサも普及をみせており、さらに、複数のセンサや映像ディスプレイをつなぐために不可欠なネットワークも発達してきている。一方、現時点では、人々が公共空間においてその場に関する情報を入手しようとする場合、キーワードによる検索をしたり能動的に行動したりしなければならないことが多い。これからは、利用者の IT リテラシが低くても、あるいは受動的であったり無意識的であったりしても、豊かな情報を享受できる自然な環境が求められている。本稿では、カメラセンサにより得られる映像ディスプレイ前の人の位置移動に基づいてリアルタイムで状況に意味づけをし、刺激コンテンツを即応的に選択提示することで人々にタイミングよく作用や影響を及ぼす機構の実現に向けた状況即応型デジタルサイネージインタフェースを提案し、その構成法について明らかにした。

A composition method of situation conformity digital-signage interface using human position and movement

Tamio KIHARA[†] Hiroshi WATANABE[†]

The number of digital signage in public spaces like as the station or the commercial establishment is increasing. Various sensors that catch situations of places are widespread. Networks indispensable to connect sensors and image displays are developing. On the other hand, people should behave actively to acquire information on the place in a public space. A natural environment that people can enjoy rich information is requested even if the people's information retrieval ability is low or people are passive or unconscious. This paper proposes the situation conformity digital-signage interface using human position and movement, and clarifies the composition method.

1. はじめに

最近では、駅や商業施設などの公共空間に設置されるデジタルサイネージの数が増加している。場の状況を捉えるセンサも普及をみせており、さらに、複数のセンサや映像ディスプレイをつなぐために不可欠なネットワークも発達してきている。一方、現時点では、人々が公共空間においてその場に関する情報を入手しようとする場合、キーワードによる検索をしたり能動的に行動したりしなければならないことが多い。これからは、日常生活の場において、利用者の IT リテラシが低くても、あるいは受動的であったり、無意識、無計画的であったりしても、豊かな情報を享受できる自然な環境が求められている。

本稿では、センサにより得られる映像ディスプレイ前の人の位置移動に基づいてリアルタイムで状況に意味づけをし、刺激コンテンツを即応的に選択提示することで人々にタイミングよく作用や影響を及ぼす機構

の実現に向けた状況即応型デジタルサイネージインタフェースを提案し、その構成法について明らかにする。

2. 関連研究

公共空間での設置を前提としたインタラクティブなディスプレイが幾つか研究開発されている¹⁾。これまで多くの場合、情報を得るためには利用者の積極的な操作が必要であり、利用者の状況によってコンテンツの内容を適応的に変化させることはなかった。

最近では、人の姿勢や動きに基づいて積極的に提示情報を変化させ、それに対する利用者のノンバーバルな反応を計測し、これらの提示と反応の組から人の心的状態を探るプロアクティブ・インタラクションモデル²⁾が提案されている。視線情報から利用者の興味や度合いを推定しコンテンツの提示形態を変化させる試みもされてきている³⁾。

これらの研究では、ある場に対する利用者の進入・退出や、その場における利用者の個々位置移動などから状況に積極的に意味づけし、それに基づいてタイミングよく映像提示することによって作用影響を及ぼす

[†]日本電信電話株式会社 NTT サイバーソリューション研究所
NTT Cyber Solutions Laboratories, Nippon Telegraph and Telephone Corporation

試みは行われていない。

3. 人の位置移動によるインタフェース

筆者らはこれまで、人の身振りや位置などの身体性の利用に着目し、複数の人が同時に利用可能なインタフェースに関する研究を進めてきた。ここでは、筆者らのこれまでの取り組みを示す。

3.1 人々の位置に基づく情報アクセス

複数の映像表示画面によって囲まれた映像空間において、人々の位置や動きなどを解析して指示命令として解釈し、表示映像を適切に変更するための情報アクセス手法について検討を行った⁴⁾。人々が映像空間内のどの位置に注目しているかを表す値として注目ポテンシャルの概念を導入し、映像空間内における人々の位置、動きの方向、複数の人の間の距離などから注目ポテンシャルを導出することを提案した。また、複数の人の指示命令が競合している場合の処理方法について考察した。

3.2 Interium

Interium⁵⁾は、複数の人が複数の映像オブジェクトを同時に操作できる情報ナビゲーションシステムである。複数の大型スクリーンを用いて構成した三面鏡型の利用空間を備えており、その内側に入った複数の人は、それぞれ映像に向かって指し示しを行うことによって映像オブジェクトを操作することができる。

3.3 みらいチューブ

みらいチューブ⁷⁾は、公共空間において複数の人が同時にインタラクション可能な巨大インタラクティブメディアシステムである。半円筒形状の空間の壁面に横長映像を投影し、認識エリア内における複数の人を複数の台のビデオカメラを用いて捕捉した。人々の位置、移動速度、加速度などを認識し、これを用いて映像コンテンツの表示内容を変化させた。図1は広告提示手法の例であり、人が興味のある映像に近づいたり、立ち止まったりすると、該当する映像を拡大して表示することを試みた。



図1. 広告提示手法

3.4 スポット情報ナビゲーション

スポット情報ナビゲーション^{9) 10)}は、実空間の状況を映像センサやRFIDを用いてリアルタイムに収集、分析し、逐次的にメタデータを付与し、その場にいる人々に対して状況に応じた情報を、予め記述したルールにより提示するシステムである。スポットとして売場や店舗を想定し、複数のスポットを相互に連携させ、人々の動線経路により情報の内容を変え、映像コンテンツによる誘導を試みた。

4. 解決すべき課題

公共空間におけるデジタルサイネージによって、人々に豊かな情報を享受してもらうためには、現状では以下のような課題がある。

- (1) デジタルサイネージで流される映像は、その場のリアルな商品やサービスと関係がないものが多い。
- (2) いわゆる広告映像が繰り返し流れているだけで、情報提示のタイミングが考慮されていない。
- (3) 人々の属性や状況に最適な情報が選択されていない。その場の環境の変化や、人間の多寡などの状況変化に追従できていない。
- (4) マーケットが、その場に即応した情報を提示させたくても、スケジュールされた映像を簡単に差し替えることができない。
- (5) 個人に適応した情報を提示することができるシステムでも、映像ディスプレイの前の利用者が携帯端末をタッチしたり、QRコードを読み取ったりする積極的な行動が求められ、利用者の負担が大きい。
- (6) 空間に対して映像ディスプレイの数が少ないとき、映像ディスプレイの前における多人数や複数の商品の所在を考慮した情報提示ができていない。
- (7) 人と映像ディスプレイの距離や、騒音レベルなどの空間特性を考慮した設計や、情報の表現ができていない。
- (8) 人が移動する動線や経路を考慮した情報提示に対応できていない。

5. アプローチと必要機能

今回、これらの課題に対し、以下のアプローチをとって解決を図ることとした。

- (1) 「場（スポット）」と「物、サービス」と「情報、刺激」をセットで管理し、このなかでの最適化を図る。
- (2) 人の人数、位置、移動方向、速度、滞留時間によるノンバーバルインタフェースを実現し、人への負担を少なくする。
- (3) 人の映像ディスプレイへの接近によってタイミン

グよく映像刺激を提示してアテンションを高め、誘導、訴求を行う。

(4) 人の間尺（ヒューマンスケール）にあわせ、音量や文字の大きさや情報量の調節をする。

ただし、今回の提案では、前提条件として以下を設定した。

- ・人への刺激は映像コンテンツのみとし、その選択はプレイリストを選択することで行う。
- ・特別な端末装置を人は持たない。
- ・人は匿名的に扱い、個人を識別同定しない。
- ・履歴の活用や学習による刺激の選択は、今回は行わない。
- ・刺激提示の自動化よりも、場や人に影響を及ぼしたい人（マーケッタ等）の意図が的確に記述できることを優先する。

アプローチの詳細を以下に示す。

5.1 フィードバックループの構成

従来のセンサネットワークの研究では、場の状況を捉えることに熱心に取り組まれてきたが、ここでは、場の状況を捉えるだけでなく、積極的にアクションを起こして、人々に作用や影響を及ぼすことを考える。図2にフィードバックループの考え方を示す。本提案では、この即応的な循環を形成することを目指す。

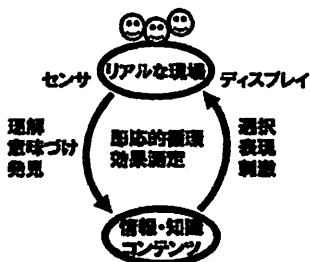


図2. フィードバックループ

5.2 情報提示のタイミング

映像ディスプレイに、例えば長さが5分程度の広告映像を繰り返し流しても、人々の関心を得るのは容易ではない。例えば、映画の予告編なら最初からみたいという要求はあるだろう。常に刺激的な映像や音声を流し続けるのではなく、タイミングよく静から動への情報提示を行えば、認知率が上がると予想できる。そこで、通常は比較的静かな背景コンテンツ（バックグラウンドプレイリスト）を流しておき、例えば人が近寄ってきたタイミングでポンというアテンション音と共に短い刺激コンテンツ（訴求プレイリスト）を提示し、訴求が終わったら背景コンテンツに戻るという手法を適用する（図3）。

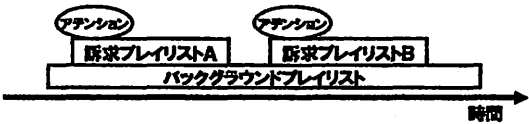


図3. 映像提示のタイミング

5.3 ノンバーバルインタフェース

ここで、スポットの人々の状況に対し、積極的に意味づけてこれを利用することを考える。

意味付与の例を以下に示す（図4）。

- ・映像ディスプレイに近寄ってきた＝何からの関心がある⇒商品購入を訴求する
- ・佇む＝興味がありしかも時間に余裕がある⇒詳しく利用法などを提示して関心を深める
- ・遠くを通りすぎようとしている＝関心がなく目的が他にある⇒お得な情報などで強く呼びかける
- ・離れる＝何かの完了⇒目先を変える、もう一押しの情報を出す
- ・映像ディスプレイ前の人が多い＝販売の好機⇒タイムセールを始める

これらを組み合わせ、タイミングよく臨機応変に畳み掛けるように情報提示することで、認知を上げたり購買意欲を刺激したりすることが可能となると考えられる。

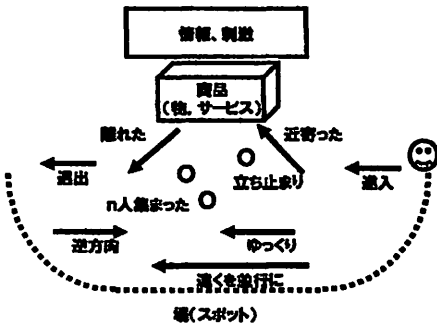


図4. 状況への意味付与の例

5.4 ヒューマンスケールのデザイン

人の間尺にあわせ、表現や音量を変えることを考える。例えば、環境の騒音レベルが高ければ、相対的に音量を大きくするなどの工夫を行う。また、1台の映像ディスプレイで同時に多人数に対して情報提示を行うことを前提に、人々の位置関係や移動方向や、人数や滞留時間によって情報量を調節することを考える。

5.5 必要な機能

これらのアプローチに従い、抽出された必要な機能を以下に示す。

- (1)センサデータの取得 映像ディスプレイ前の人々

の位置、移動方向、移動速度、対人距離、人数、滞留時間、商品に手を伸ばしたかなどをリアルタイムで把握するためのセンサデータ取得機能が必要である。

(2)状況への意味づけ リアルタイムに得られたセンサデータに対し、それを状況として積極的に意味付与(メタデータ化)する必要がある。

(3)刺激の結び付け 状況の意味付与に対し、それにふさわしいタイミングでの販売促進用の映像プレイリストを結びつける必要がある。

(4)条件判定と選択 状況の意味付与に対し条件判定の閾値を設定したり、競合が起きた場合の選択優先順位の指定ができたりする必要がある。

(5)刺激の再生 刺激コンテンツを即時再生する必要がある。このとき、次々と起こる状況の変化によって、現在の刺激コンテンツの再生完了後に次の再生をするのか、中断して再生をするのかを指定できる必要がある。また、音量調節などの指定ができる必要がある。

これらのフィードバックループの実現に加えて、以下の機能も必要となる。

(6)記述エディタ スポットの様々な状況に対して、どのような刺激コンテンツをどのタイミングで提示するかは、マーケッタが決めるものとする。このマーケッタの意図を正しく記述するためのエディタが必要となる。また、記述内容を保存し、再利用できる必要がある。試行錯誤しやすいように簡易に記述編集でき、リアルタイムで結果が反映されることが望ましい。

(7)ガイドラインとノウハウ このような構成のスポットを実空間に設置するためのガイドラインや、ノウハウが必要となる。

6. 提案手法

6.1 システムの構成

これらの必要機能を実現するための手法の提案内容と、それを実現するシステムの構成を以下に示す。

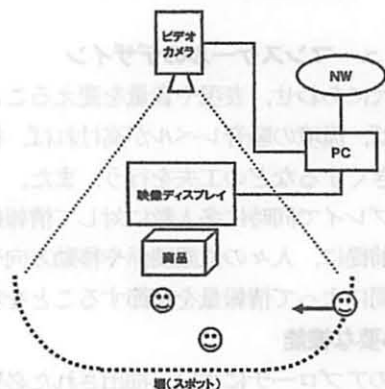


図5. スポットの構成

1つのスポットは、映像ディスプレイとこれに映像信号を出力するPCとを中心に構成される(図5)。天井にはビデオカメラがあり、この映像ディスプレイ前の様子を撮像して利用する。これをカメラセンサとして利用するPCは、映像信号を出力するPCと同一で構わない。映像ディスプレイの前に、訴求したい商品があることを原則とする。

カメラセンサは、人がいない状態を背景とし、この背景との差分を随時処理することにより人影を抽出し、この人影にオブジェクトIDを付与することで利用する。人影処理の例を図6~8に示す。



図6. カメラセンサの画像

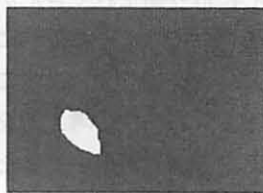


図7. 人影の抽出

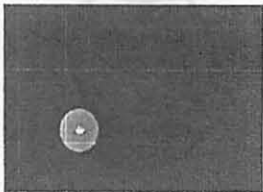


図8. 論理空間での人影オブジェクト

ここで、状況と映像プレイリストを簡単に結びつけるために、人々の位置や移動に対してラベルを付与することが重要であると考え、スポットをセルと呼ぶ領域に分割してこれを記述に利用することにした。セルによるスポットの分割を図9に示す。

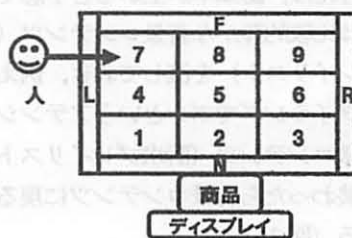


図9. セルによるスポットの分割

ここでセルは、映像ディスプレイの前の場所を9分割したもので、近い／中くらい／遠いと、進入してきた／映像ディスプレイの正面／退出していくという3×3のマトリクスとなっている。これらに便宜的に1～9の数字を割り当てラベルとして利用する。これら9個のセルの枠外には右側R、左側L、近いN、遠いFのエリアがあり、枠外と枠内のセルとの移動を記述できるようにする。これにより、セルのラベルを使って、正規表現により条件照合することができる。

例えば、

52 中央から映像ディスプレイ前への移動
[456][123] 映像ディスプレイへ近い場所への移動
L[147] 左側からの進入

のように移動を技術し、条件判定に利用する。セルとセルの境界は、スライドさせて調整することができるようにする。スポットの映像ディスプレイの大きさや通路の大きさ等によって近い／遠い感覚が異なってくるためである。

6.2 ソフトウェアの構成

今回の提案手法を実現するために作成したソフトウェアの構成を図10に示す。

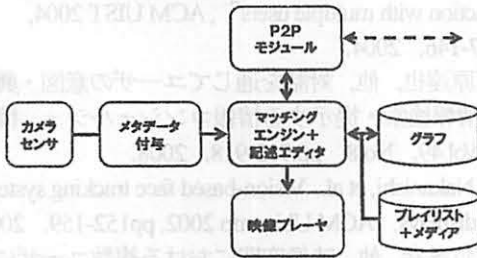


図10. ソフトウェアの構成

カメラセンサからのデータはメタデータ付与機能によって意味づけられる。メタデータと映像プレイリストを結びつけ条件設定したデータフロログラムによりマッチングエンジンが駆動し、映像プレーヤによって逐次映像が再生される。このデータフロログラムをグラフと呼び、作成編集するエディタを提供した(図11)。グラフはコンポーネントを結線したものであり、各コンポーネントには、カメラセンサやプレイリストや映像プレーヤをハンドリングするもの、条件設定や優先順位をするセレクトなどがある。各コンポーネントのプロパティを修正することで、柔軟な設定が可能になっている。グラフはXMLのファイルとして保存され活用される。

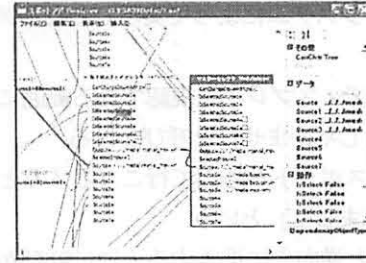


図11. グラフとプロパティ編集の例

映像ディスプレイへの映像の再生では、複数の短い刺激コンテンツを連続して提示する必要から、映像と映像の間に短い黒画面が挟まらないように、予め映像コンテンツファイルをバッファリングしておき、瞬時に切り替えられるようにした。

また、映像ディスプレイとカメラセンサを1対1で結びつけるだけではなく、隣のスポットの情報も利用するために、複数スポット間のメタデータの相互参照が可能なP2Pモジュールを実装した。

7. 試行システム

今回、提案手法の有効性を確認する試行システムを構築した。映像ディスプレイ前の商品を映像により訴求することを目的に、サンプル映像により実際に人々に対して情報提示することを行った。試行システムの構成を図12に示す。動作の様子を図13に示す。



図12. 試行システムの構成

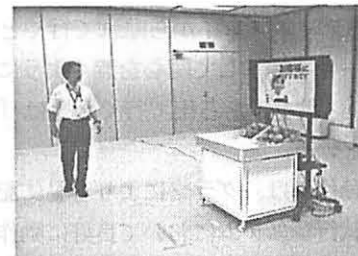


図13. 試行システムによる動作の様子

今回の試行システムによって実現した動作例を以下に示す。

- (1) 人が映像ディスプレイを視認できる範囲に進入すると「いらっしゃいませ」と呼びかける。
- (2) 人がそのスポットから出て行こうとすると「ありがとうございました」という。
- (3) 順路方向と逆向きに進入する人には呼びかけない。
- (4) 映像ディスプレイ前に来ると商品を推薦する。
- (5) 6秒以上立ち止まっていると商品の詳しい説明（レシピ等）に切り替わる。
- (6) 映像ディスプレイ前の右側の商品に手を伸ばすとその商品情報が出る。
- (7) 映像ディスプレイ前の左側の商品に手を伸ばすとその商品情報が出る。
- (8) 映像ディスプレイ前に2人の人がいると1人とは違う商品を推薦する。
- (9) 映像ディスプレイ前の2人の人の間の距離が近いとまとめ買いを推奨する。
- (10) 映像ディスプレイから遠いところを人が通り過ぎようとする「ちょっと待っていただけますか」と呼びとめる。
- (11) そのスポットの人数が3人以上だとタイムセールを始める情報を出す。
- (12) 人数が5人以上になると群集状態になったとしてその場の商品訴求をやめる。

次に、映像ディスプレイと人との距離によって、商品（ケーキ）を訴求する情報表現が変化する動作例を以下に示す。

- (1) 人が映像ディスプレイから遠いと大きな音声で「ケーキはいかがですか」と呼びかける。映像はケーキのイメージ映像で文字は含まれない。
- (2) 人が映像ディスプレイに近づいてくると、ケーキの種類の数やそれらの名前がわかる程度の見出し文字を映像として提示する。
- (3) 人が映像ディスプレイに最も近づくと、そのケーキの精細な動画映像や詳細なテキスト情報などを提示する。
- (4) 人が映像ディスプレイから離れて行くときは、違う種類のケーキを提示して関心を惹くようにする。
- (5) 人がそのスポットから出て行くときは「またお願いします」という。

これらの動作例は、グラフにより容易に記述することができ、試行システムにおいて良好に動作した。セルのラベルを利用した条件設定によって柔軟な設定変更ができた。カメラセンサの画角とセルの境界を調整することによって、人がちょうど映像ディスプレイの

内容を視認できるような距離になった時にタイミングよく刺激映像を提示することができた。しかしながら、人数が多くなり、人影の交差などが起きると動作が不安定になるという問題がわかった。この試行システムでは、人影の位置移動のみを利用し、例えば性別や子供かどうかによって情報を出しわけたりはしていない。

8. おわりに

本稿では、デジタルサイネージによる情報利用のインタフェースの関連研究について述べ、人の位置移動によるインタフェースに関する筆者らのこれまでの取り組みを整理し、これらを踏まえて、センサにより得られる映像ディスプレイ前の人の位置移動に基づいてリアルタイムで状況に意味づけをし、刺激コンテンツを即応的に選択提示することで人々にタイミングよく作用や影響を及ぼす機構の実現に向けた状況即応型デジタルサイネージインタフェースを提案し、その構成法について明らかにした。今後、実験を通して提案手法の有効性を検証していく。

参考文献

- 1) D. Vogel, et al., "Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users", ACM UIST 2004, pp137-146, 2004.
- 2) 河原達也, 他, 対話を通じてユーザの意図・興味を探り情報検索・提示する情報コンシェルジュ, 情報処理, Vol.49, No.8, pp.912-918, 2008.
- 3) Y. Nakanishi, et al., Vision-based face tracking system for large displays, ACM Ubicomp 2002, pp152-159, 2002.
- 4) 大竹孝幸, 他, 映像空間における複数ユーザによる情報アクセス方法, 情報研報, Vol.2000, No.18, pp.25-30, 2000.
- 5) 木原民雄, 実写映像の多人数操作による情報ナビゲーション, DICO2002 シンポジウム論文集, pp.9-12, 2002.
- 6) 木原民雄, 他, 場のアートとネットワークのアート, 情報処理学会論文誌, Vol.44, No.2, pp.212-221, 2003.
- 7) 篠原章夫, 他, 公共の場でのインタラクティブメディア実証実験「みらいチューブ」実験報告, 情報研報, Vol.2006, No.14, pp.163-168, 2006.
- 8) 篠原章夫, 他, 公共空間における巨大インタラクティブメディアシステム: みらいチューブ, ヒューマンインタフェース学会論文誌, Vol.9, No.3, pp.305-312, 2007.
- 9) 田辺弘実, 他, 実空間メタデータ収集に基づく情報ナビゲーション, DICO2003 シンポジウム論文集, pp.273-276, 2003.
- 10) 田辺弘実, 他, スポット情報ナビゲーション技術, NTT 技術ジャーナル, Vol.15, No.18, pp.37-40, 2003.