

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-179285
(P2007-179285A)

(43) 公開日 平成19年7月12日(2007.7.12)

(51) Int. Cl. F I テーマコード (参考)
G06T 7/00 (2006.01) G06T 7/00 300D 5L096
 G06T 7/00 350C

審査請求 未請求 請求項の数 9 O L (全 16 頁)

(21) 出願番号	特願2005-376676 (P2005-376676)	(71) 出願人	504174135 国立大学法人九州工業大学 福岡県北九州市戸畑区仙水町1番1号
(22) 出願日	平成17年12月27日(2005.12.27)	(74) 代理人	110000154 特許業務法人はるか国際特許事務所
		(72) 発明者	山川 烈 福岡県北九州市若松区ひびきの2-4 国立大学法人九州工業大学内
		(72) 発明者	堀尾 恵一 福岡県北九州市若松区ひびきの2-4 国立大学法人九州工業大学内
		(72) 発明者	相川 暁 福岡県北九州市八幡西区千代ヶ崎1-15-31 コーポせいざん107号
		Fターム(参考)	5L096 BA17 FA34 HA09 HA11 JA09 JA11

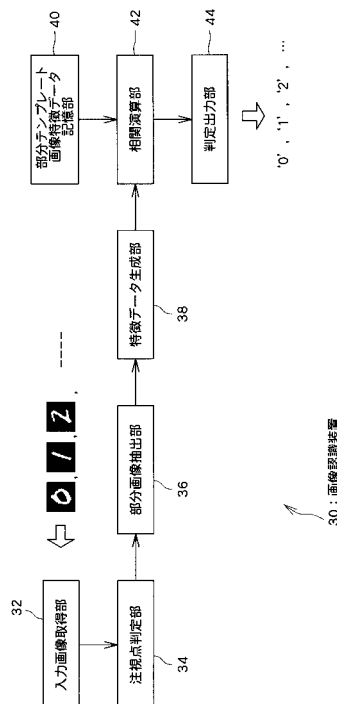
(54) 【発明の名称】 画像認識装置、画像認識方法及びプログラム

(57) 【要約】

【課題】複雑な調整を要せず、しかも短時間で入力画像に示される記述内容の認識を行うこと。

【解決手段】入力画像を取得する入力画像取得部32と、入力画像における所定数の注視点を判定する注視点判定部34と、入力画像における前記各注視点の位置の部分画像に関するデータ(特徴データ)を生成する特徴データ生成部38と、複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータ(特徴データ)を記憶する部分テンプレート画像特徴データ記憶部40と、を含み、入力画像に係る特徴データと、前記各記述内容に対応づけて前記部分テンプレート画像特徴データ記憶部40に記憶される特徴データと、に基づいて、入力画像に示される記述内容を判定する。

【選択図】 図18



【特許請求の範囲】**【請求項 1】**

入力画像を取得する入力画像取得手段と、

前記入力画像における所定数の注視点を判定する注視点判定手段と、

前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成手段と、

複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段と、

前記部分入力画像データ生成手段により生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定手段と、

を含むことを特徴とする画像認識装置。

【請求項 2】

請求項 1 に記載の画像認識装置において、

前記注視点判定手段は、前記所定数のユニットが配置された競合層を備える自己組織化マップに、前記入力画像における記述箇所の位置座標のベクトルを入力するとともに、前記競合層の前記所定数のユニットの重みベクトルを前記所定数の注視点の位置座標のベクトルとして取得する、

ことを特徴とする画像認識装置。

【請求項 3】

請求項 2 に記載の画像認識装置において、

前記部分テンプレート画像データ記憶手段は、前記各記述内容を示すテンプレート画像の前記所定数の注視点の位置の部分画像のデータのそれぞれを、前記所定数のユニットのいずれかに対応づけて記憶し、

前記記述内容判定手段は、前記各記述内容について、相互に対応関係にあるユニットに対応づけられた、前記入力画像に係るデータと前記各テンプレート画像に係るデータとの各相関量を算出するとともに、それら相関量に基づいて前記入力画像に示される記述内容を判定する、

ことを特徴とする画像認識装置。

【請求項 4】

請求項 3 に記載の画像認識装置において、

前記自己組織化マップの前記競合層には、前記所定数のユニットが回転対称性を有するようにして配置され、

前記記述内容判定手段は、前記回転対称性に従って前記対応関係を変更しながら、前記入力画像に係るデータと前記各テンプレート画像に係るデータとの相関量を算出する、

ことを特徴とする画像認識装置。

【請求項 5】

請求項 1 乃至 4 のいずれかに記載の画像認識装置において、

前記部分入力画像データ生成手段は、前記入力画像における前記各注視点の位置の部分画像を前記入力画像から抽出するとともに、それら部分画像に基づいて、それらの特徴を示すデータを生成する、

ことを特徴とする画像認識装置。

【請求項 6】

請求項 5 に記載の画像認識装置において、

前記部分入力画像データ生成手段は、前記入力画像における前記各注視点の位置の部分画像の特徴を示すデータとして、該部分画像に示される記述内容の一部の延伸方向を示すデータを生成する、

ことを特徴とする画像認識装置。

【請求項 7】

10

20

30

40

50

請求項 6 に記載の画像認識装置において、

前記部分入力画像データ生成手段は、適応部分空間自己組織化マップに前記入力画像における前記各注視点の位置の部分画像を入力し、前記適応部分空間自己組織化マップの応答を、該部分画像に示される記述内容の一部の延伸方向を示すデータとして取得する、ことを特徴とする画像認識装置。

【請求項 8】

入力画像を取得する入力画像取得ステップと、

前記入力画像における所定数の注視点を判定する注視点判定ステップと、

前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成ステップと、

複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段を参照して、前記部分入力画像データ生成ステップで生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定ステップと、

を含むことを特徴とする画像認識方法。

【請求項 9】

入力画像を取得する入力画像取得手段、

前記入力画像における所定数の注視点を判定する注視点判定手段、

前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成手段、

複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段、及び

前記部分入力画像データ生成手段により生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定手段

としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は画像認識装置、画像認識方法及びプログラムに関し、特に入力画像からそこに示される記述内容を判定する画像認識装置、画像認識方法及びプログラムに関する。

【背景技術】

【0002】

従来、パターン認識の手法は数多く提案されているが、その中で下記非特許文献 1 に開示されたネオコグニトロンは、入力画像の局所的な特徴を階層型ニューラルネットワークの各層で抽出するとともに、それらを段階的に統合して、最終的な入力画像のパターン認識を行うものであって、位置ずれ、拡大、縮小に強い画像認識を可能とする技術として知られている。

【非特許文献 1】福島邦彦著，「神経回路網と情報処理」，朝倉書店，1989年

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、上記のネオコグニトロンを構成する階層型ニューラルネットワークは非常に多くのパラメータを含んでおり、その調整が非常に困難である。さらに、ネオコグニトロンには、画像認識処理に膨大な時間を要するという問題がある。

【0004】

本発明は上記課題に鑑みてなされたものであって、その目的は、複雑な調整を要せず、

10

20

30

40

50

しかも短時間で入力画像に示される記述内容の認識が可能な画像認識装置、画像認識方法及びプログラムを提供することにある。

【課題を解決するための手段】

【0005】

上記課題を解決するために、本発明に係る画像認識装置は、入力画像を取得する入力画像取得手段と、前記入力画像における所定数の注視点を判定する注視点判定手段と、前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成手段と、複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段と、前記部分入力画像データ生成手段により生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定手段と、を含むことを特徴とする。

10

【0006】

また、本発明に係る画像認識方法は、入力画像を取得する入力画像取得ステップと、前記入力画像における所定数の注視点を判定する注視点判定ステップと、前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成ステップと、複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段を参照して、前記部分入力画像データ生成ステップで生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定ステップと、を含むことを特徴とする。

20

【0007】

さらに、本発明に係るプログラムは、入力画像を取得する入力画像取得手段、前記入力画像における所定数の注視点を判定する注視点判定手段、前記入力画像における前記各注視点の位置の部分画像に関するデータを生成する部分入力画像データ生成手段、複数の記述内容のそれぞれに対応づけて、該記述内容を示すテンプレート画像における前記所定数の注視点の位置の部分画像のそれぞれに関するデータを記憶する部分テンプレート画像データ記憶手段、及び前記部分入力画像データ生成手段により生成されるデータと、前記各記述内容に対応づけて前記部分テンプレート画像データ記憶手段に記憶されるデータと、に基づいて、前記入力画像に示される記述内容を判定する記述内容判定手段としてコンピュータを機能させるためのプログラムである。このプログラムは、例えばCD-ROMやDVD-ROM等のコンピュータ読み取り可能な情報記憶媒体に格納されてもよい。

30

【0008】

本発明では、複数の記述内容（例えば文字や数字など）のそれぞれに対応づけて、該記述内容を示すテンプレート画像の注視点の位置の部分画像に関するデータが、予め記憶手段に記憶される。このデータは、部分画像そのものを示すデータであってもよいし、部分画像の特徴等を示すデータであってもよい。また、入力画像における所定数の注視点が判定されるとともに、その位置の部分画像に関するデータが生成される。このデータも、部分画像そのものを示すデータであってもよいし、部分画像の特徴等を示すデータであってもよい。そして、入力画像から生成されるデータと、各記述内容のテンプレート画像について予め記憶手段に記憶されるデータと、に基づいて、入力画像に示される記述内容が判定される。本発明によれば、入力画像における所定数の注視点の判定、入力画像の部分画像に関するデータの生成、該データとテンプレート画像の部分画像に関するデータとの照査等、比較的調整が容易で、しかも負荷の軽い処理により、入力画像に示される記述内容を判定でき、短時間で入力画像に示される記述内容の認識ができる。

40

【0009】

本発明の一態様では、前記注視点判定手段は、前記所定数のユニットが配置された競合層を備える自己組織化マップに、前記入力画像における記述箇所的位置座標のベクトルを

50

入力するとともに、前記競合層の前記所定数のユニットの重みベクトルを前記所定数の注視点の位置座標のベクトルとして取得する。自己組織化マップ (Self Organizing Map; SOM) は、トポロジカルマッピングを自己組織的に獲得する工学モデルであり、入力ベクトルを受け付ける入力層のユニットと、入力層のユニットと重みベクトル (ユニットを特徴付けるパラメータ) を介して結合される競合層のユニットと、を含んで構成される。入力ベクトルを受け付けられると、該入力ベクトルに最も近い値を有する重みベクトルを介して結合されている競合層のユニット、及びその近傍のユニットについて、それらの重みベクトルが入力ベクトルに近づくよう更新される。この自己組織化マップに入力画像における記述箇所の位置座標を入力することで、入力画像の記述箇所の位置座標を、競合層の各ユニットの重みベクトルとして得ることができる。

10

【0010】

この態様においては、前記部分テンプレート画像データ記憶手段は、前記各記述内容を示すテンプレート画像の前記所定数の注視点の位置の部分画像のデータのそれぞれを、前記所定数のユニットのいずれかに対応づけて記憶してよい。また、前記記述内容判定手段は、前記各記述内容について、相互に対応関係にあるユニットに対応づけられた、前記入力画像に係るデータと前記各テンプレート画像に係るデータとの各相関量を算出するとともに、それら相関量に基づいて前記入力画像に示される記述内容を判定するようにしてよい。

【0011】

ここで、前記自己組織化マップの前記競合層には、前記所定数のユニットが回転対称性を有するようにして配置されてよい。この場合、前記記述内容判定手段は、前記回転対称性に従って前記対応関係を変更しながら、前記入力画像に係るデータと前記各テンプレート画像に係るデータとの相関量を算出してもよい。こうすれば、入力画像に記述内容が回転した状態で示されていても、適切に記述内容の判定が行えるようになる。

20

【0012】

また、本発明の一態様では、前記部分入力画像データ生成手段は、前記入力画像における前記各注視点の位置の部分画像を前記入力画像から抽出するとともに、それら部分画像に基づいて、それらの特徴を示すデータを生成する。こうすれば、入力画像に示される記述内容の認識をさらに高精度に行うことができる。

【0013】

この態様では、前記部分入力画像データ生成手段は、前記入力画像における前記各注視点の位置の部分画像の特徴を示すデータとして、該部分画像に示される記述内容の一部の延伸方向を示すデータを生成するようにしてよい。前記部分入力画像データ生成手段は、適応部分空間自己組織化マップに前記入力画像における前記各注視点の位置の部分画像を入力し、前記適応部分空間自己組織化マップの応答を、該部分画像に示される記述内容の一部の延伸方向を示すデータとして取得するようにしてよい。

30

【0014】

この態様では、前記部分入力画像データ生成手段は、適応部分空間自己組織化マップに前記入力画像における前記各注視点の位置の部分画像を入力し、前記適応部分空間自己組織化マップの応答を、該部分画像の特徴を示すデータとして取得するようにしてよい。適応部分空間自己組織化マップ (Adaptive Subspace Self Organizing Map; ASSOM) は、様々な方向を持つ二次元正弦波画像を事前に学習しておくことで、方位選択性細胞のような機能を獲得できることが知られているもので、この適応部分空間自己組織化マップを用いることで、部分画像に示された記述内容の一部の特徴、すなわち記述内容の一部の延伸方向を示すデータを容易に生成できるようになる。

40

【発明を実施するための最良の形態】**【0015】**

以下、本発明の実施形態について図面に基づき詳細に説明する。

【0016】

図1は、本発明の実施形態に係る画像認識方法を示すフロー図である。本画像認識方法

50

では、手書きされた「0」、「1」、「2」、「3」、「4」、「5」、「6」、「7」、「8」、「9」のいずれかの数字（記述内容）を示す入力画像を取得すると、この入力画像に示された数字が上記の0～9のうちいずれであるかを判定する。ここでは、簡単のため、入力画像に示された手書き数字を判定する方法を説明するが、手書き文字、手書きの絵柄、印刷数字、印刷文字、印刷絵柄等、入力画像に示された様々な種類の記述内容を判定するのに本発明を適用してよいのはもちろんである。

【0017】

図1に示すように、本画像認識方法では、まず手書きの数字が示された入力画像を取得する（S101）。入力画像は、例えば記述部分の画素に0よりも大きな画素値が関連づけられ、それ以外に0の画素値が関連づけられたものであり、二値画像、グレースケール画像、フルカラー画像等、あらゆる形式の画像を用いることができる。

10

【0018】

そして、この入力画像の中の所定数（ここでは19個）の注視点の位置を判定する（S102）。図2は、この注視点の判定処理に用いられる自己組織化マップの構成を示す図である。自己組織化マップ12（SOM）はニューラルネットワークの一種であって、コンピュータ上で仮想的に実現されるものであり、2個のユニットが配置された入力層14と、注視点の個数と同数である19個のユニットが配置された競合層16から構成されている。図3は、図2に示す自己組織化マップの学習処理を示すフロー図である。同図に示すように、この学習処理では、まず競合層16の各ユニットjの結合重みベクトル $w^{(j)}$ を初期化する（ $j = 1 \sim 19$ ；S201）。初期状態においては、図14に示すようにして、競合層16に19個のユニットjが回転対称を有するようにして配置される。次に、入力層14のユニットに、入力画像10における画素値が0よりも大きい画素（記述箇所の画素）の位置座標 $v_i = (x_i, y_i)$ を入力ベクトルとして入力する（S202）。競合層16の各ユニットjは結合重みベクトル $w^{(j)}$ が関連づけられており（ $j = 1 \sim 19$ ）、次式（1）に従って、入力ベクトル v_i に最も近い値の結合重みベクトルwに関連づけられたユニットの番号 c_i を勝者ユニットの番号として決定する（S203）。ここで、tは結合重みベクトルwの更新時刻を示す。

20

【0019】

【数1】

$$c_i = \arg \min_j \|v_i - w^{(j)}(t)\| \quad \dots (1)$$

30

【0020】

そして、入力画像10における輝度が0より大きい全ての画素の位置座標について、S202及びS203の処理を実行して、各入力ベクトル v_i に対する勝者ユニット c_i を決定する（ $i = 1, 2, 3, \dots$ ）。その後、次式（2）に従って、競合層16の各ユニットjの重みベクトル $w^{(j)}$ を更新する（S204）。ここで、 $h_{j c_i}(t)$ は、近傍係数であり、例えば時刻tにおけるユニットjと勝者ユニット c_i との距離に関する減少関数として定義される。

40

【0021】

【数2】

$$w^{(j)}(t+1) = \frac{\sum_i h_{j c_i}(t) v_i}{\sum_i h_{j c_i}(t)} \quad \dots (2)$$

【0022】

図5は、図2に示される入力画像10から得られる入力ベクトルを全て入力し、学習が完了した後の結合重みベクトル $w^{(j)}$ の配置を示している。同図に示すように、競合層16の各ユニットjの結合重みベクトル $w^{(j)}$ は、入力画像10における画素値が0よ

50

りも大きい画素の位置座標に近づいており、本画像認識方法では、これらの結合重みベクトル $w^{(j)}$ を、入力画像 10 の 19 個の注視点の位置座標として利用する。なお、図 6 は、手書きの 0 ~ 9 の数字が示された各入力画像を用いた学習後における、重みベクトル $w^{(j)}$ を示している。

【0023】

図 1 に戻り、本画像認識方法では、次に入力画像の 19 個の注視点の位置の各周辺画像を抽出する (S103)。図 7 は、図 5 に示される重みベクトル $w^{(j)}$ の位置を中心とした 7 画素 × 7 画素の大きさの各部分画像を示している。

【0024】

次に、こうして得られる入力画像の 19 個の部分画像のそれぞれについて、特徴データを生成する (S104)。特徴データは 100 個 (= 10 × 10) の数値データから構成されており、図 8 には、各部分画像の特徴データが画像として示されている。同図に示される特徴データの画像は、部分画像に示される線分 (数字の一部) の方向が共通している場合に (上下左右にシフトしていても)、近似した模様を示すことが分かる。このように、特徴データは、各部分画像に示される記述内容の方向を示すデータとなっている。ここでは、こうした特徴データを、適応部分空間自己組織化マップ (Adaptive Sub-space Self Organizing Map; ASSOM) を用いて生成しており、その処理については後に詳述する。なお、各部分画像の特徴データは、例えばガボールフィルタを用いて生成してもよい。

【0025】

図 1 に再び戻り、以上の処理は、事前に、テンプレート画像群に対しても同様に行われる。すなわち、本画像認識方法では、図 9 に示すような、複数組の手書き数字を示す画像をテンプレート画像として用意する (S105)。そして、各テンプレート画像に対して、S101 の処理と同様にして、19 個の注視点の位置座標を算出する (S106)。さらに、各テンプレート画像から、そのテンプレート画像の 19 個の注視点の位置から 7 画素 × 7 画素の大きさの部分画像を抽出する (S107)。そして、各部分画像の特徴データを、S104 の処理と同様にして生成する (S108)。こうして生成される各部分画像の特徴データは、元の入力画像に示されている記述内容 (0, 1, 2, 3, ...) に関連づけて、事前にハードディスク記憶装置や EEPROM 等の記憶手段に記憶される (S109)。図 10 には、「0」の手書き数字が示されたテンプレート画像から生成された 19 個の特徴データの画像を示している。

【0026】

本画像認識方法では、次に、S104 で得られた入力画像の j 番目の部分画像 (自己組織化マップの競合層 16 のユニット j に関する部分画像) の特徴データ U_j と、S109 で記憶された、いずれかのテンプレート画像の j 番目の部分画像の特徴データ V_j と、の相関値 R_j を次式 (4) に従って算出する。ここで、特徴データ U_j は、 $u_{j1} \sim u_{jM}$ から構成され、特徴データ V_j は、 $v_{j1} \sim v_{jM}$ から構成される ($M = 100$)。そして、入力画像とテンプレート画像との相関値 S を、次式 (5) に従って算出し、同様にして、他の全てのテンプレート画像と入力画像との相関値 S を算出する。

【0027】

【数 3】

$$R_j = \frac{\langle U_j, V_j \rangle}{\|U_j\| \cdot \|V_j\|} = \frac{\sum_{m=1}^M u_{jm} v_{jm}}{\sqrt{\sum_{m=1}^M u_{jm}^2} \sqrt{\sum_{m=1}^M v_{jm}^2}} \quad \dots (3)$$

$$S = \frac{\sum_{j=1}^n R_j}{N} \quad \dots (4)$$

10

20

30

40

50

【0028】

さらに、上記のように、入力画像やテンプレート画像の注視点を判定するのに用いる自己組織化マップは、その競合層に19個のユニットが回転対称性を有するようにして配置されているので、この回転対称性に従って、入力画像の部分画像とテンプレート画像の部分画像との対応関係を変更しながら、上記相関値 R_j 及び相関値 S を算出する(S110)。その後、最も大きな値の相関値 S に対応するテンプレート画像の特徴データ群に関連づけて事前に記憶されている記述内容(数字)を、入力画像の記述内容として決定し、それを出力する(S111)。

【0029】

ここで、各部分画像の特徴データを生成する処理について説明する。図11は、部分画像の特徴データを生成するのに用いられる、適応部分空間自己組織化マップ(ASSOM)を示す図である。同図に示す適応部分空間自己組織化マップ20は、部分画像の画素数である49個(=7×7)のユニットが配置された入力層22と、10×10個のユニットが配置された競合層24と、から構成されている。競合層24の各ユニット j には、図12に模式的に示されるような部分空間26が関連づけられており、この部分空間26は49次元の互いに直交する基底ベクトル $b_h^{(j)}$ により張られている($h < 49$)。適応部分空間自己組織化マップ20の学習のためには、5枚の2次元正弦波画像からなる学習用入力画像組が多数用意される。各学習用入力画像組は、図13に示すように、1枚の2次元正弦波画像28a、それを上方向に所定画素(ここでは2画素)だけシフトしてなる上シフト画像28b、左方向に所定画素(ここでは2画素)だけシフトしてなる左シフト画像28c、右方向に所定画素(ここでは2画素)だけシフトしてなる右シフト画像28d、及び下方向に所定画素(ここでは2画素)だけシフトしてなる下シフト画像28eの5枚の画像からなる。

【0030】

適応部分空間自己組織化マップ20の学習に用いられる入力ベクトル $x(t_p)$ は、各学習用入力画像組の中で t_p 番目の画像の各画素値を要素とする、49(=7×7)次元のベクトルである。図14に示すように、学習時には、まず競合層24上の各ユニット j の部分空間26の基底ベクトル $b_h^{(j)}$ を乱数により初期化するとともに、この際、基底ベクトル $b_1^{(j)}$ と基底ベクトル $b_2^{(j)}$ とを正規直交化する(S301)。そして、学習用入力画像組から得られる各入力ベクトル $x(t_p)$ に対して、次式(5)により、競合層24上の各ユニット j に関連づけられた部分空間26への射影ベクトル $X^{(j)}(t_p)$ のノルムを計算する(S302)($j = 1 \sim 100$, $t_p = 1 \sim 5$ (図15参照))。さらに、次式(6)により、代表勝者ユニット c を決定する(S303)。

【0031】

【数4】

$$\|X^{(j)}(t_p)\| = \sqrt{\sum_h (x^T(t_p)b_h^{(j)})^2} \quad \dots (5)$$

$$c = \arg \max_j \left\{ \sum_{t_p} \|X^{(j)}(t_p)\|^2 \right\} \quad \dots (6)$$

【0032】

その後、代表勝者ユニット c と、競合層24においてその近傍に配置されたユニットの基底ベクトルを、次式(7)に従って更新する(S304)。ここで (t) は学習係数であり、 $h_c^{(j)}(t)$ は近傍係数である。これにより、代表勝者ユニット c と、その近傍のユニットの基底ベクトル $b_h^{(j)}$ は、例えば学習用入力画像組に係る入力ベクトル群の平均ベクトル等に近づけられる($h = 1, 2$)。また、基底ベクトル $b_h^{(j)}$ の更新時には、それらの正規直交化が行われる。そして、上記S302乃至S304の処理を、学習係数及び近傍範囲を小さくしながら所定回数だけ繰り返し、さらに同様の処理を他の学習用入力画像組についても実行する。

10

20

30

40

50

【 0 0 3 3 】

【 数 5 】

$$b_h^{(j)}(t+1) = \left[I + \lambda(t) h_c^{(j)}(t) \frac{x(t_p) x^T(t_p)}{\|X^{(j)}(t_p)\| \|x(t_p)\|} \right] b_h^{(j)}(t) \quad \dots (7)$$

【 0 0 3 4 】

図 1 6 は、学習の結果として得られた、競合層 2 4 の各ユニット j の部分空間 2 6 の基底ベクトル $b_h^{(j)}$ を示している ($j = 1 \sim 100$)。同図において、各円は競合層 2 4 のユニットを示しており、その中に示された 2 つの画像は、該ユニットの部分空間 2 6 の 2 つの基底ベクトル $b_1^{(j)}$ 及び $b_2^{(j)}$ を示している。このようにして学習がなされた適応部分空間自己組織化マップ 2 0 においては、次式 (8) のようにして、競合層 2 4 のユニット j の部分空間 2 6 への射影ベクトル $X^{(j)}$ のノルムを、入力ベクトル x で除した値である、競合層 2 4 のユニット j の応答 u_j を計算すると、上下左右に並行移動した画像については、ほぼ同じ箇所の競合層 2 4 のユニットの応答が大きくなる。

10

【 0 0 3 5 】

【 数 6 】

$$u_j = \frac{\|X^{(j)}\|}{x} = \frac{\sqrt{\sum_h (x^T b_h^{(j)})^2}}{\|x\|} \quad \dots (8)$$

20

【 0 0 3 6 】

例えば、図 1 7 (a) に示すようにして、幅が 1 画素のラインが描かれた入力画像 (同図最上段)、幅が 3 画素のラインが描かれた入力画像 (同図 2 段目)、幅が 5 画素のラインが描かれた入力画像 (同図 3 段目)、幅が 1 画素であり最上段に比して左に 2 画素だけシフトした入力画像 (同図 4 段目)、及び幅が 1 画素であり最上段に比して上に 2 画素だけシフトした入力画像 (同図 5 段目) を用意し、それらの画素値を要素とする入力ベクトルを、学習後の適応部分空間自己組織化マップ 2 0 に入力し、その応答を求めると、同図 (b) のようになる。同図 (b) に表された各画像は、同図 (a) の対応する箇所 30 の入力画像の画素値を要素とする入力ベクトルの応答を示しており、その各画素値は上記応答 u_j となっている。この図から分かるように、学習済みの適応部分空間自己組織化マップ 2 0 によれば、位相ズレの関係にある入力画像に対して、ほぼ同様の応答をする。

30

【 0 0 3 7 】

本実施形態に係る画像認識方法では、入力画像の各注視点の位置の部分画像を抽出すると、この部分画像の画素値を要素とするベクトルを、上述の学習を済ませた適応部分空間自己組織化マップ 2 0 に入力し、その応答 ($u_1 \sim u_{100}$) を特徴データとして用いている。これにより、部分画像に表された記述部分の方向を示す特徴データを得て、それを記述内容の判定に用いることができ、認識精度を向上させることができる。

【 0 0 3 8 】

また、本画像認識方法によれば、入力画像の注視点を複数判定して、その局所的特徴を基にして記述内容の判定を行うようにしているが、これにより人間の視覚系に近い高精度のパターン認識を実現できる。また、自己組織化マップ 1 2 の処理、適応部分空間自己組織化マップ 2 0 の処理等、比較的負荷の軽い処理により、またパラメータ調整の少ない処理により、簡易且つ短時間で入力画像に示される記述内容の認識ができる。

40

【 0 0 3 9 】

そして、上述の画像認識方法を用いる画像認識装置は、例えば以下のようにして実現される。図 1 8 は、本実施形態に係る画像認識装置 3 0 の機能ブロック図である。画像認識装置 3 0 はコンピュータを中心として構成されるものであり、所定の認識プログラムを実行することにより、同図に示される各機能が実現される。なお、認識プログラムはコンピ

50

ユーザ読み取り可能な情報記憶媒体に格納され、そこからコンピュータにインストールされてもよいし、インターネット等のデータ通信ネットワークからコンピュータにダウンロードされてもよい。

【0040】

同図に示すように、画像認識装置30は、機能的には入力画像取得部32、注視点判定部34、部分画像抽出部36、特徴データ生成部38、部分テンプレート画像特徴データ記憶部40、相関演算部42及び判定出力部44を含んでいる。ここで、入力画像取得部32は、入力画像を取得するものであり、例えばデジタルカメラのような撮像手段や、各種方式のタッチパネルにより構成することができる。注視点判定部34は、図1のS102の処理を行うものであり、入力画像における所定数の注視点を判定する。すなわち、所定数のユニットが配置された競合層16を備える自己組織化マップ12に、入力画像における記述箇所の位置座標のベクトルを入力するとともに、競合層16の前記所定数のユニットの結合重みベクトルを前記所定数の注視点の位置座標のベクトルとして取得する。

10

【0041】

部分画像抽出部36は、図1のS103の処理を行うものであり、入力画像における前記各注視点の位置の部分画像を抽出するものである。また、特徴データ生成部38は、S104の処理を行うものであり、各部分画像の特徴データを生成するものである。ここでは、適応部分空間自己組織化マップ20に部分画像のベクトルを入力して、その応答を特徴データとして取得する。この特徴データは、各部分画像に示される記述内容の一部の延伸方向を示すものとなる。

20

【0042】

次に、部分テンプレート画像特徴データ記憶部40は、図1のS109の処理結果として得られるデータを記憶するものであり、複数の記述内容(0~9)のそれぞれに対応づけて、該記述内容を示すテンプレート画像における注視点の位置の部分画像のそれぞれに関する特徴データを記憶する。そして、相関演算部42は、図1のS110の処理を行うものであり、特徴データ生成部38により生成される、入力画像に関する特徴データ群と、部分テンプレート画像特徴データ記憶部40に各記述内容に関連づけて記憶される特徴データ群と、の相関値Sを算出する。判定出力部44は、図1のS111の処理を行うものであり、こうして算出される相関値Sの最大値を与える特徴データ群に関連づけられた記述内容を、入力画像の記述内容として、例えばモニタにより表示出力する。

30

【0043】

以上の画像認識装置30によれば、非常に軽い処理で入力画像に示された記述内容の認識が可能であり、例えば携帯情報端末等の非力なコンピュータを用いても、高精度の画像認識が可能となる。

【図面の簡単な説明】

【0044】

【図1】本発明の実施形態に係る画像認識方法を示すフロー図である。

【図2】注視点の判定に用いる自己組織化マップを示す図である。

【図3】自己組織化マップによる学習処理を示すフロー図である。

【図4】自己組織化マップの結合重みベクトルの初期配置を示す図である。

40

【図5】自己組織化マップの学習後の結合重みベクトルの配置を示す図である。

【図6】様々な入力画像に対する、自己組織化マップの学習後の結合重みベクトルの配置を示す図である。

【図7】入力画像から各注視点の位置の部分画像を抽出した様子を示す図である。

【図8】各部分画像の特徴データの画像を示す図である。

【図9】テンプレート画像例を示す図である。

【図10】各テンプレート画像に対して予め生成される特徴データの画像を示す図である。

。

【図11】特徴データの生成に用いられる適応部分空間自己組織化マップを示す図である。

。

50

【図12】適応部分空間自己組織化マップの競合層のユニットに関連づけられた部分空間を示す図である。

【図13】適応部分空間自己組織化マップの学習に用いる学習用入力画像組の一例を示す図である。

【図14】適応部分空間自己組織化マップの学習処理を示すフロー図である。

【図15】入力ベクトルが競合層の各ユニットの部分空間に射影される様子を示す図である。

【図16】学習済みの適応部分空間自己組織化マップの競合層の各ユニットの基底ベクトルを示す図である。

【図17】適応部分空間自己組織化マップの応答例を示す図である。

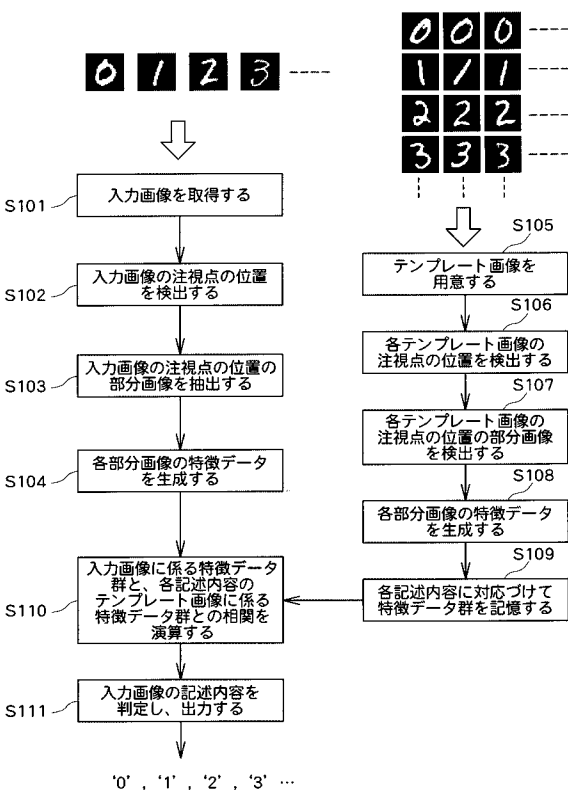
【図18】本発明の実施形態に係る画像認識装置の機能ブロック図である。

【符号の説明】

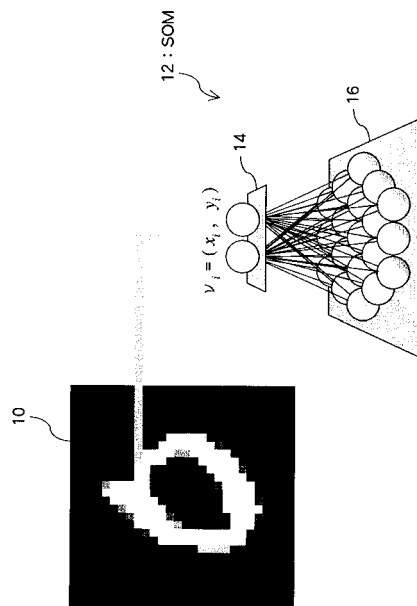
【0045】

10 入力画像、12 自己組織化マップ(SOM)、14, 22 入力層、16, 24 競合層、20 適応部分空間自己組織化マップ(ASSOM)、26 部分空間、28 学習用入力画像組、30 画像認識装置、32 入力画像取得部、34 注視点判定部、36 部分画像抽出部、38 特徴データ生成部、40 部分テンプレート画像特徴データ記憶部、42 相関演算部、44 判定出力部。

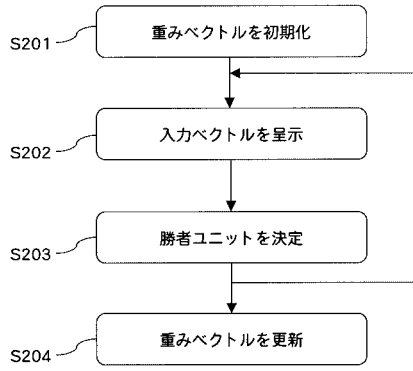
【図1】



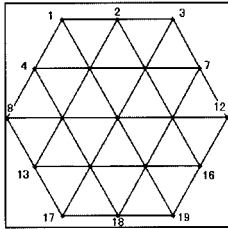
【図2】



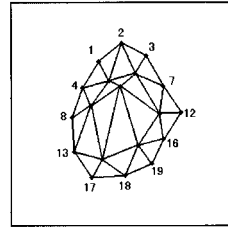
【図3】



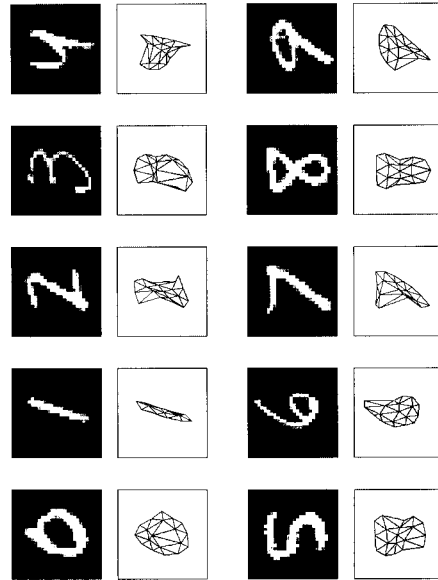
【図4】



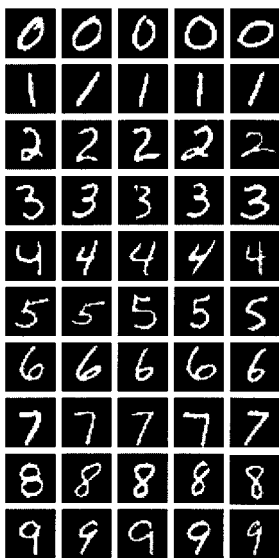
【図5】



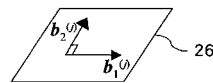
【図6】



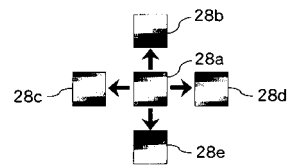
【図9】



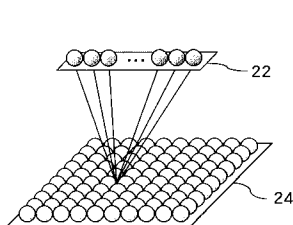
【図12】



【図13】

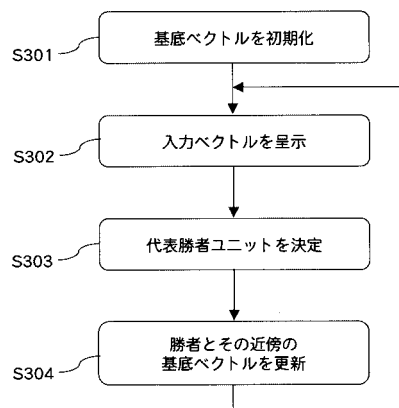


【図11】

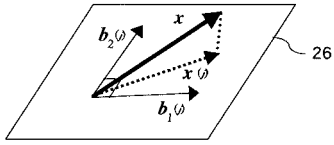


20: ASSOM

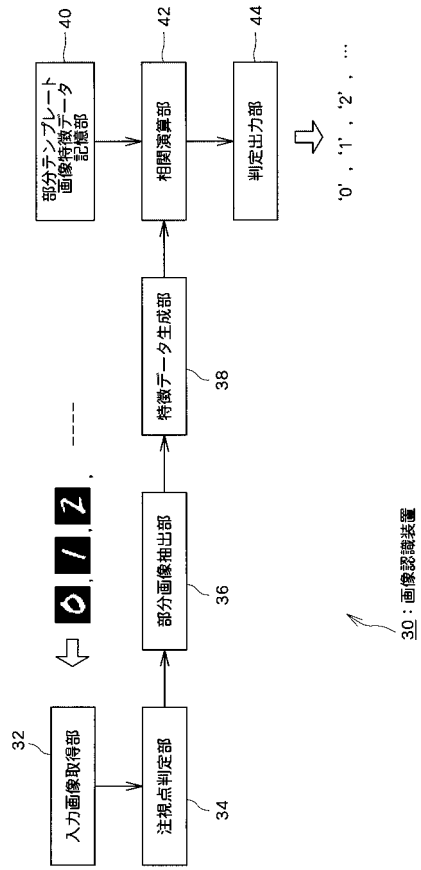
【図14】



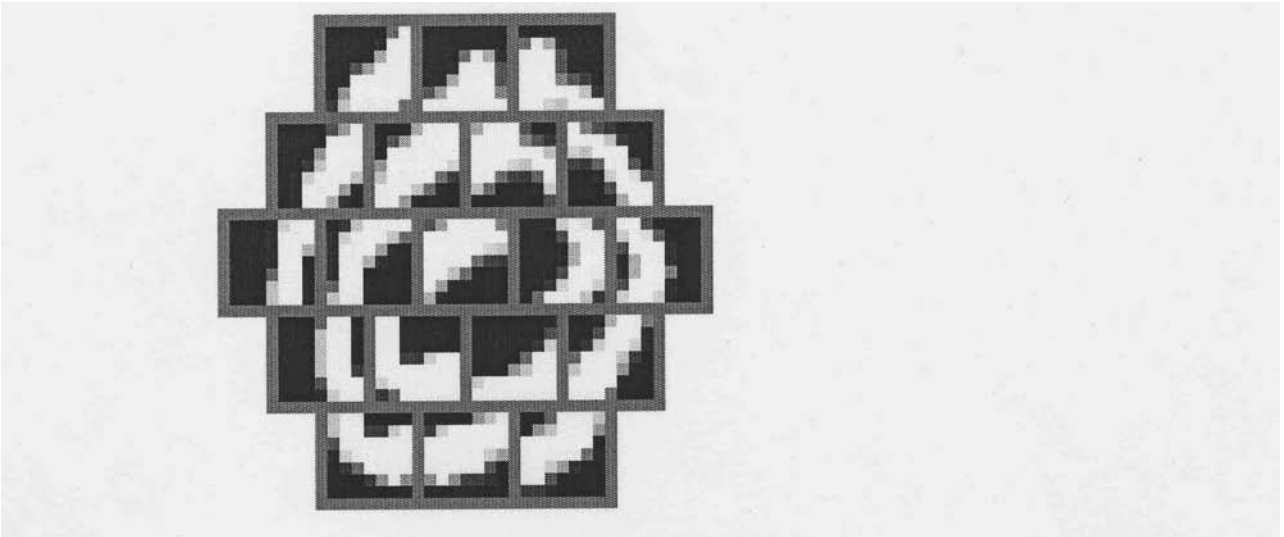
【図 15】



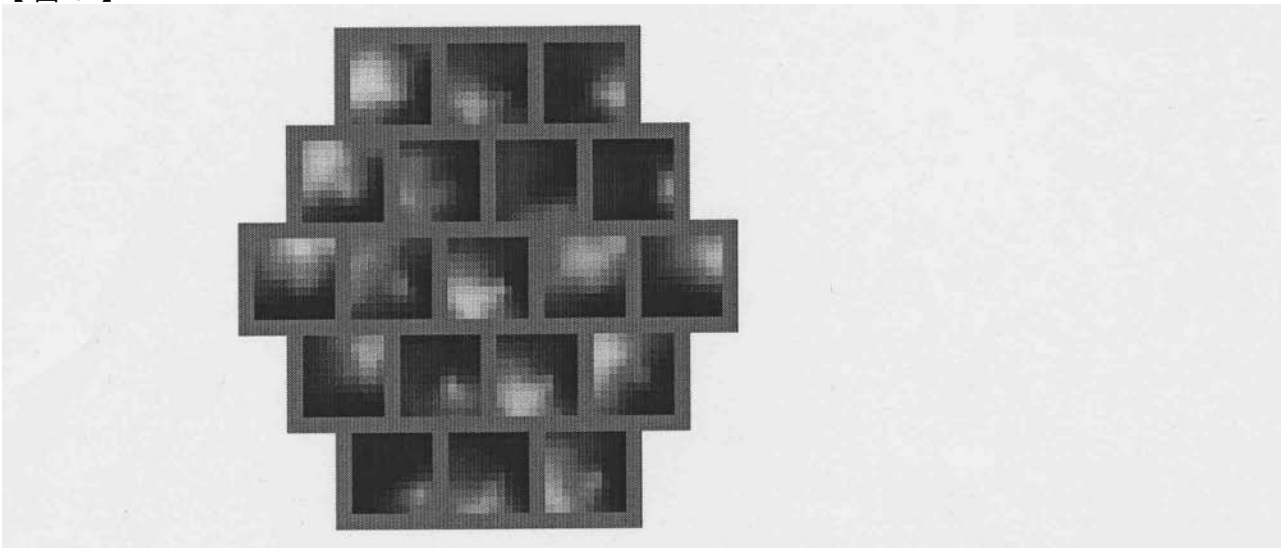
【図 18】



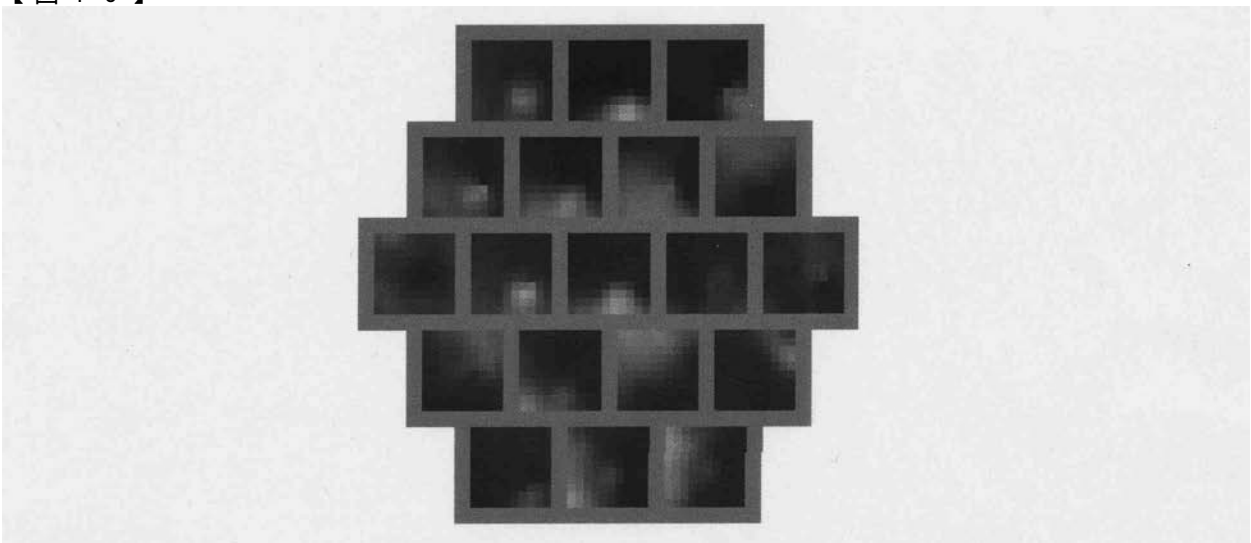
【図 7】



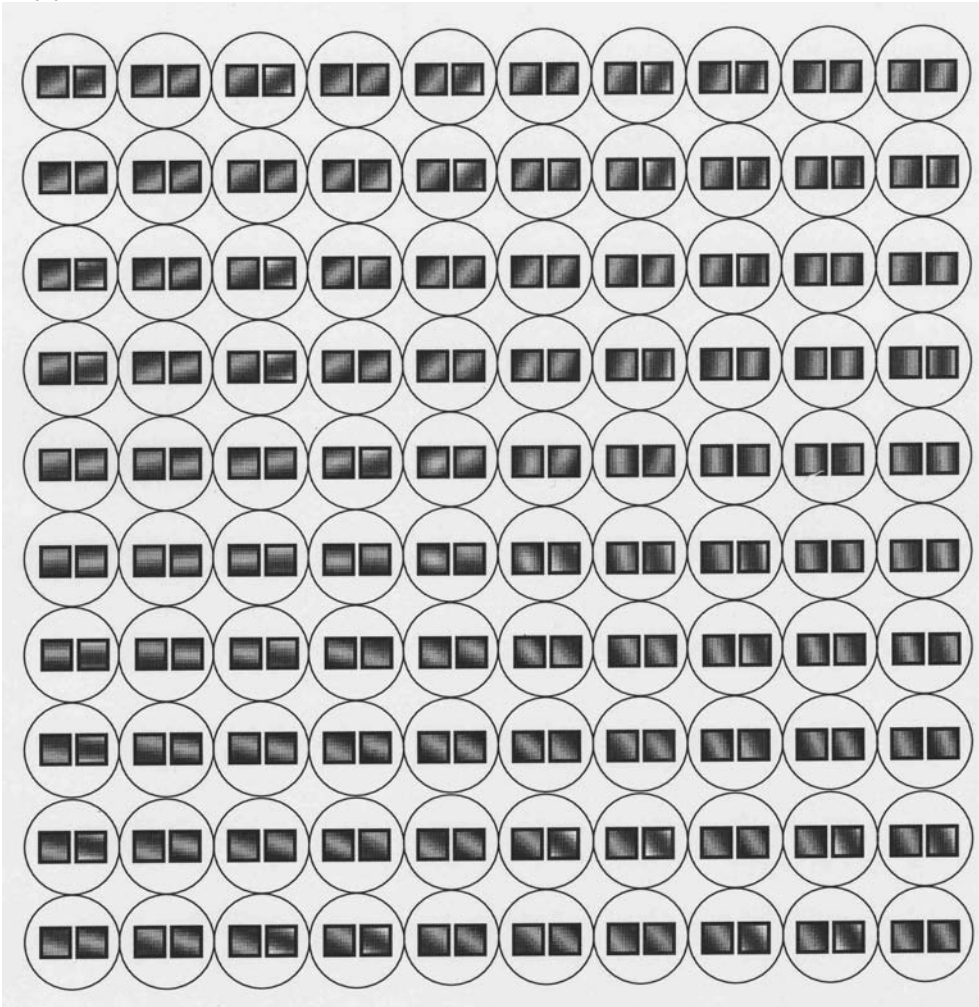
【 図 8 】



【 図 10 】



【 図 16 】



【 図 17 】

