(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第5506722号
(P5506722)
(45)発行日　平成26年5月28日(2014.5.28)　(24)登録日　平成26年3月28日(2014.3.28)

(51)Int.Cl.　　F I
G06T 7/00 (2006.01)　　G06T 7/00 350B
G06N 3/00 (2006.01)　　G06N 3/00 560C

請求項の数 15　外国語出願　(全 14 頁)

(21)出願番号　特願2011-51823(P2011-51823)
(22)出願日　平成23年3月9日(2011.3.9)
(65)公開番号　特開2011-210252(P2011-210252A)
(43)公開日　平成23年10月20日(2011.10.20)
審査請求日　平成26年1月21日(2014.1.21)
(31)優先権主張番号　12/732,225
(32)優先日　平成22年3月26日(2010.3.26)
(33)優先権主張国　米国(US)

早期審査対象出願

(73)特許権者　597067574
ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウエイ　201
201 BROADWAY, CAMBRIDGE, MASSACHUSETTS 02139, U.S.A.
(74)代理人　100110423
弁理士　曾我　道治
(74)代理人　100094695
弁理士　鈴木　憲七
(74)代理人　100111648
弁理士　梶並　順

最終頁に続く

(54)【発明の名称】マルチクラス分類器をトレーニングするための方法

(57)【特許請求の範囲】
【請求項１】
マルチクラス分類器をトレーニングするための方法であって、
前記分類器によって求められるメンバーシップ確率に基づいて、１組のアクティブ画像から１つの問い合わせ画像を選択するステップであって、前記アクティブ画像は、ラベル無しである、選択するステップと、
前記問い合わせ画像の前記メンバーシップ確率に基づいて、１組のトレーニング画像から１つのサンプル画像を選択するステップであって、前記トレーニング画像は、ラベル付きである、選択するステップと、
前記問い合わせ画像及び前記サンプル画像を、出力デバイスにおいてユーザーに対して表示するステップと、
入力デバイスを用いてユーザーからの応答を得るステップであって、前記応答は、イエス照合又はノー照合である、得るステップと、
前記イエス照合が得られた場合には、前記サンプル画像のラベルと共に前記問い合わせ画像を前記トレーニングセットに追加し、そうでない場合には、前記マルチクラス分類器を得るために、所定の数のノー照合に達するまで、前記選択するステップ、前記表示するステップ及び前記得るステップを繰り返すステップと、
１組のラベル無し画像に前記マルチクラス分類器を適用するステップであって、１組の検出結果を得る、適用するステップと、
前記１組の検出結果のメンバーシップ確率を求めるステップと、

所定のしきい値未満の前記メンバーシップ確率を有する前記１組の検出結果を前記１組のアクティブ画像に関連付けるステップと、
前記マルチクラス分類器をトレーニングし直すステップであって、前記１組の検出結果を精緻化する、トレーニングし直すステップと、を含み、
各ステップは、プロセッサによって実行される
マルチクラス分類器をトレーニングするための方法。
【請求項２】
一対一サポートベクトルマシン、及び逐次最小最適化を用いて、前記メンバーシップ確率を推定するステップをさらに含む
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項３】
前記メンバーシップ確率は、ロジスティック回帰分類器を用いて推定される請求項２に記載のマルチクラス分類器をトレーニングするための方法。
【請求項４】
前記メンバーシップ確率は、複数のクラスにわたってフィッシャー判別分析を用いて推定される請求項２に記載のマルチクラス分類器をトレーニングするための方法。
【請求項５】
前記所定の数のノー照合に達した場合には、新たなクラスラベルを前記問い合わせ画像に割り当てるステップをさらに含む
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項６】
前記最初の１組のトレーニング画像は画像のシードセットからランダムに選択される請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項７】
前記最初の１組のトレーニング画像は画像サーチエンジンのデータベースに問い合わせることによって選択される請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項８】
前記最初の１組のトレーニング画像はユーザーによって供給される請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項９】
複数の応答を得るために複数のユーザーが用いられる請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１０】
前記問い合わせ画像を選択するステップは、予想される情報価値（ＥＶＩ）に基づき、目的関数が予想されるリスクとユーザーラベル付けのコストを結合し、
前記サンプル画像を選択するステップは、前記問い合わせ画像との一致の最大尤度に基づく
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１１】
凝集クラスタリングによって、トレーニングされたクラスをマージするステップをさらに含む
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１２】
クラス間類似度スコアが前記凝集クラスタリング内の所定のしきい値よりも大きい場合に、２つのクラスを結合するステップをさらに含む
請求項１１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１３】
２つのクラスが前記ユーザーに対して表示され、前記２つのクラスがマージされるべきであるか否かを尋ねるとき、前記応答が前記イエス照合である場合に、前記２つのクラス

を結合するステップをさらに含む
請求項１１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１４】
用途及び分野の仕様に基づいて、予想されるリスクの要因を調整するステップをさらに含む
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【請求項１５】
前記ユーザーから無視応答を得るステップであって、前記１組のトレーニング画像から前記問い合わせ画像を削除する、得るステップをさらに含む
請求項１に記載のマルチクラス分類器をトレーニングするための方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、包括的にはマルチクラス分類器をトレーニングすることに関し、より詳細には、能動的選択及び２値フィードバックを用いてトレーニングすることに関する。
【背景技術】
【０００２】
マルチクラス画像分類器をトレーニングするには、多数のラベル無し画像を必要とする。多数の画像をインテリジェントに利用するのは、難しい問題である。能動トレーニング（能動学習と呼ばれる場合もある）は、情報を提供する画像を選択し、２値及びマルチクラス分類のために、分類器をトレーニングすることを目標とする。マルチクラス能動トレーニング法が、必要とされるトレーニング画像の数を削減することに成功する場合であっても、それらの方法は、以下の理由のために、ユーザーインタラクションの観点から大きな労力を要する可能性がある。
（ｉ）ラベル付けのために問い合わせられるラベル無し画像毎に、ユーザーは、多数のクラスをふるい分けして、正確なクラスを入力しなければならない。特に画像の場合、この形において入力を与えることは難しい可能性があり、莫大な数（又は未知数）のクラスが存在するときには、不可能であることもある。
（ｉｉ）クラスの数が増えるほど、長い時間がかかり、大きな労力を要する。
（ｉｉｉ）ラベル付けの際に、ユーザーインタラクションは、間違いに陥りやすい。
（ｉｖ）ラベル付けの際に、全てのユーザーが一致する必要があるので、分散型のラベル付けには馴染みにくい。
【発明の概要】
【発明が解決しようとする課題】
【０００３】
画像のデータベースは、サイズ、画像の種類ともに、益々増加しつつある。何千もの画像クラスを有するのが一般的である。規模が大きくなっても実用的である方法を設計するために、ユーザーのためのラベル付け及びインタラクションのより容易な形態を可能にすることが不可欠である。
【課題を解決するための手段】
【０００４】
本発明の実施の形態は、ユーザーからのイエス／ノーフィードバックのみを必要とするマルチクラス分類器をトレーニングするための方法を提供する。インタラクションの各インスタンス中に、ユーザーは２つの画像を提示され、それらの画像が同じクラスに属するか否かを指示しなければならない。そのような入力を与えることは極めて容易であり、常に２つの画像しか照合する必要がないので、人的間違いに陥りにくくもある。また、その方法は、複数のユーザーによる分散型のラベル付けも可能にする。
【０００５】
本発明は、２値フィードバックモデルにおける、予想される情報価値（ＥＶＩ：Ｅｘｐｅｃｔｅｄ　Ｖａｌｕｅ　Ｉｎｆｏｒｍａｔｉｏｎ）に基づく能動的選択手順を提供する

。
【０００６】
その手順は、ユーザーがラベル付けする労力を最小限に抑えながらも、精度を最大限にするように、ラベル付けのための画像対を繰り返し選択する。
【０００７】
その方法は、１００クラスもの数の分類問題に関して、従来のトレーニングモデルに比べて、ユーザー管理を実質的に最小限に抑えることができる。
【０００８】
本発明者らが知る限り、これは、一人又は複数のユーザーからのイエス／ノー入力応答のみを必要とする初めてのマルチクラス能動トレーニング手法である。
【発明の効果】
【０００９】
本発明では、ユーザーからの２値フィードバックのみを必要とする新たなマルチクラス能動トレーニングの枠組みを記述する。本発明の方法は、ユーザートレーニング時間及び労力を大きく削減する。その方法は、母集団不均衡及び雑音のような現実世界の問題に対してロバストである。
【図面の簡単な説明】
【００１０】
【図１】本発明の実施の形態による、マルチクラス分類器をトレーニングするための方法のブロック図である。
【図２】図１の方法のための擬似コードのためのブロック図である。
【図３】本発明の実施の形態による、２値フィードバックを用いるインタラクティブトレーニングの概略図である。
【図４】従来技術のマルチクラスインタラクティブトレーニングの概略図である。
【発明を実施するための形態】
【００１１】
図１は、本発明の実施の形態によるマルチクラス分類器１５０を能動的にトレーニングするための方法を示す。図２は、比較可能な擬似コードを示す。その方法は、当該技術分野において知られているように、メモリ、入力／出力インターフェースを含む、プロセッサ１００において実行することができる。
【００１２】
アクティブセット１０１は、多数のラベル無し画像を含み、本発明の方法は、それらの画像から、一人又は複数のユーザーが入力／出力デバイス１４１及び１４２を用いて照合する（１４０）するための問い合わせ画像１０２を選択する（１２０）。マルチクラス分類器をトレーニングするために、１組のラベル付きトレーニング画像１０４も用いられる。それらのトレーニング画像は、画像サーチエンジン１０６のデータベース１０５に問い合わせることによって得ることができるか、又はユーザーが供給することができる。
【００１３】
基礎的な分類器として、一対一サポートベクトルマシン（ＳＶＭ：Ｓｕｐｐｏｒｔ　Ｖｅｃｔｏｒ　Ｍａｃｈｉｎｅｓ）が用いられる。マルチクラスの場合、一対一ＳＶＭが用いられる（クラス対毎に分類器がトレーニングされる）。
【００１４】
従来のマルチクラス能動トレーニングでは、ユーザーラベル付けのために、ラベル無し問い合わせ画像を選択する必要がある。しかしながら、本発明によれば、ユーザー入力は、２値（イエス／ノー）のみであるので、本発明では、照合用として既知のクラスからのラベル付き画像もユーザーに対して表示する必要がある。トレーニングセット１０４からサンプル画像１０３を選択することは、本発明の枠組みが必要とする能動トレーニングの新たな態様である。
【００１５】
本発明は、問い合わせ選択１２０及びサンプル選択１３０に焦点を当てる。

【００１６】
本発明による問い合わせ及びサンプル選択のための手法は確率的であり、すなわち、現在のトレーニングセットに基づいて、アクティブセット１０１内の画像のためのクラスメンバーシップ確率推定値が得られる。この実施の形態では、本発明は、ラベル無し画像に関するマルチクラス確率推定のための一対一ＳＶＭとのペアワイズ結合と組み合わせて、プラッツの方法を用いて、ＳＶＭマージンに基づいて２値確率を推定する。プラッツの方法は、最適化及び機械学習において逐次最小最適化（ＳＭＯ：Ｓｅｑｕｅｎｔｉａｌ Ｍｉｎｉｍａｌ Ｏｐｔｉｍｉｚａｔｉｏｎ）としてよく知られている。たとえば、米国特許第７，５９９，８９７号を参照されたい。プラッツの方法は、大きな二次計画法（ＱＰ：Ｑｕａｄｒａｔｉｃ Ｐｒｏｇｒａｍｍｉｎｇ）問題を、考えられる最小の一連のＱＰ問題に分割し、その後、それらの問題を分析的に解く。
【００１７】
代替的には、本発明では、ロジスティック曲線を用いてメンバーシップの確率を予測するために、ロジスティック回帰分類器を用いることができる。線形判別分析（ＬＤＡ：Ｌｉｎｅａｒ Ｄｉｓｃｒｉｍｉｎａｎｔ Ａｎａｌｙｓｉｓ）、たとえば、フィッシャー線形判別分析を用いて、２つ以上のクラスを特徴付ける、画像内の特徴の一次結合を求めることもできる。
【００１８】
問い合わせ選択１２０は、推定された（１１０）クラスメンバーシップ確率を用いて、アクティブセット１０１から問い合わせ画像１０２を選択する。問い合わせ画像のための推定されたメンバーシップ確率に基づいて、サンプル選択手順１３０は、トレーニングセット１０４からサンプル画像１０３を選択する。
【００１９】
出力デバイス１４１上で照合するために（１４０）、問い合わせ画像及びサンプル画像対がユーザーに対して表示される。入力デバイス１４２を用いて、「イエス照合」応答が得られる場合には、問い合わせ画像及びサンプル画像が同じクラスに属することを指示しており、この問い合わせ画像が、サンプル画像からのクラスラベルと共に、トレーニングセット１０４に追加される。
【００２０】
「ノー照合」応答が得られる場合には、異なるサンプル画像を得るために、サンプル選択手順１３０が繰返し呼び出される。後者の場合、ユーザーによって、新たなクラスラベルが開始され、問い合わせ画像に割り当てられる。多数のクラスが存在するとき、新たなクラスを開始するには、多数のユーザー応答を必要とする可能性がある。以下では、本発明者らは、クラスタマージングと共に、高速の新クラス初期化ステップを通じて、これをいかに克服するかを記述する。
【００２１】
その後、１組のラベル無し画像にマルチクラス分類器を適用して、１組の検出結果を求め、その１組の検出結果のためのメンバーシップ確率を求めることができる。その後、所定のしきい値未満のメンバーシップ確率を有する１組の検出結果を、１組のアクティブ画像と関連付けることができ、マルチクラス分類器をトレーニングし直して、その１組の検出結果を精緻化することができる。
【００２２】
そのような仕組みを通して、シード画像セットから最初にランダムに選択されたわずかなトレーニング画像を用いて、トレーニングを開始することができる。
【００２３】
その過程が継続するとき、その能動選択手順は、別の試験セットにおいて同様の分類速度を達成するのに、ランダム選択よりもはるかに少ない問い合わせしか必要としない。その方法は、利用可能である場合には、正確なクラスラベル付けに関するフィードバックも利用できることに留意されたい。しかしながら、２値フィードバックは、適用性を一般化し、新たな未知の分類問題におけるトレーニングを可能にする。

【００２４】
ユーザーにペアワイズマストリンク（ｍｕｓｔ－ｌｉｎｋ）及びキャンノットリンク（ｃａｎｎｏｔ－ｌｉｎｋ）制約を要求することによって、データクラスタリングとの関連において、２値入力が用いられてきた。しかしながら、分類とは対照的に、全てのデータがバッチクラスタリング（ｂａｔｃｈ　ｃｌｕｓｔｅｒｉｎｇ）環境を通して入手可能であるので、２値入力をそのまま分類に拡張することはできない。
【００２５】
能動トレーニング法
２値フィードバック能動トレーニングには、２つの部分がある。
（ｉ）アクティブセットから問い合わせ画像を選択すること。
（ｉｉ）既知のクラスラベルを有するサンプル画像を選択して、問い合わせ画像と共にユーザーに対して表示すること。
【００２６】
問い合わせ選択
問い合わせ選択の目的は、情報を提供する画像、すなわち、分類の精度を改善することに繋がる可能性がある画像を問い合わせることである。本発明では、問い合わせ画像との一致の最大尤度に基づく、予想される情報価値（ＥＶＩ）を用いる。ＥＶＩは、決定理論の分野においてよく知られている。
【００２７】
その概念は、目的関数に基づいてサンプルを選択することであり、その関数は、予想されるリスクと、ユーザーラベル付けのコストとを結合する。ｋクラス問題のためのリスク行列Ｍを考える。その行列内の成分M_{ij}は、真のラベルｉを有する画像をクラスｊに属するものと誤分類することに関連するリスクを指示する。正しい分類は、リスクを被ることはなく、それゆえ、行列Ｍの対角要素は０である。
【００２８】
本発明者らは、ラベル無し画像ｘのための推定されるクラスメンバーシップ分布を、次式のように表す。
【００２９】
【数１】

$$\mathbf{p}_x = \{p^1{}_x, \ldots, p^k{}_x\}$$

【００３０】
ｘのための真のクラスメンバーシップ分布は、未知であるので、実際の誤分類リスクを求めることはできない。代わりに、本発明では、ｘのための予想される誤分類リスクを次式（１）のように求める。
【００３１】
【数２】

$$\mathcal{R}_{\mathcal{L}}^{\{x\}} = \sum_{i=1}^{k}\sum_{j=1}^{k} M_{ij}\cdot(p_x^i|\mathcal{L})\cdot(p_x^j|\mathcal{L}) \qquad (1)$$

【００３２】
ただし、Ｌは、確率が推定される１組のラベル付きサンプルである。試験セットＴは、Ｎ個の画像$x_1, \ldots x_N$を有する。試験セットにわたって予想される全リスク（サイズによって正規化される）は、次式（２）のようになる。
【００３３】

【数３】

$$\mathcal{R}_{\mathcal{L}} = \frac{1}{|T|} \sum_{x \in T} \sum_{i=1}^{k} \sum_{j=1}^{k} M_{ij} \cdot (p_x^i | \mathcal{L}) \cdot (p_x^j | \mathcal{L}) \qquad (2)$$

【００３４】

上記の式は、全リスクを計算しながら、試験セットが入手できることを要求する。典型的には、試験セットは、予め入手することはできない。それゆえ、本発明では、予想されるリスクを求めるために、アクティブセットＡ１０１内の画像を用いることができる。実際には、分類に関する大部分の研究は、試験セットが存在しない場合に、代わりのデータを用いて誤分類リスクを推定する。用途及び分野の仕様によって、予想されるリスク要因を調整することができる。

【００３５】

多くのシナリオにおいて、完全に入手可能な１組のラベル無し画像がアクティブセットとして用いられ、それは典型的には非常に大きいので、アクティブセットに関するリスクの推定値は、かなり信頼性が高い。ユーザーからそのラベルを入手することによって、セットＡ内の画像ｙがラベル付きトレーニングセット１０４に追加される場合には、アクティブセットに関して予想されるリスク低減は、次式（３）のようになる。

【００３６】

【数４】

$$\mathcal{R}_{\mathcal{L}} - \mathcal{R}_{\mathcal{L}'} = \frac{1}{|\mathcal{A}|} \sum_{x \in \mathcal{A}} \sum_{i=1}^{k} \sum_{j=1}^{k} M_{ij} \cdot (p_x^i | \mathcal{L}) \cdot (p_x^j | \mathcal{L})$$

$$- \frac{1}{|\mathcal{A}'|} \sum_{x \in \mathcal{A}'} \sum_{i=1}^{k} \sum_{j=1}^{k} M_{ij} \cdot (p_x^i | \mathcal{L}') \cdot (p_x^j | \mathcal{L}') \qquad (3)$$

【００３７】

ただし、次式が成り立つ。

【００３８】

【数５】

$$L' = L \cup \{y\}, \text{ and } A' = A \setminus \{y\}$$

【００３９】

上記の式は、問い合わせ画像ｙの値を取り込み、画像ｙをラベル付きトレーニングセットに追加する。しかしながら、本発明では、ユーザーから画像ｙのためのフィードバックを得ることに関連するコストも考慮する必要がある。画像ｙに関するユーザーラベル付けを得るコストは、Ｃ（ｙ）であると仮定する。

【００４０】

本発明の方法では、誤分類のリスク低減を最大にしながら、かかったコストを低減する画像を能動的に選択することを望む。リスク低減及びラベル付けコストが、同じ単位において判定されると仮定すると、問い合わせ画像ｙのためのＥＶＩを表す結合目的関数（ｊｏｉｎｔ　ｏｂｊｅｃｔｉｖｅ）は、次式（４）のようになる。

【００４１】

【数６】

$$V(y) = \mathcal{R}_{\mathcal{L}} - \mathcal{R}_{\mathcal{L}'} - C(y) \qquad (4)$$

【００４２】

上記の式の項R_Lはｙ、すなわち、問い合わせのための選択されることになる例から独立している。それゆえ、ＥＶＩを最大にするための能動的選択は、最小化として次式（５）で表すことができる。

【００４３】

【数７】

$$y^* = \underset{y \in \mathcal{A}}{\operatorname{argmin}} \quad \mathcal{R}_{\mathcal{L}'} + C(y) \qquad (5)$$

【００４４】

上記の枠組みは、その分野に特有である、リスク及びラベル付けコストの任意の概念を利用することができる。たとえば、本発明では、或る特定のクラスに属する例を誤分類することは、他よりも費用がかかる可能性があるという事実を捕捉することができる。そのような概念は、画像が潜在的に危険性のある腫瘍を含むか否かを判断できるように医療画像を分類する場合に極めて役に立つことがある。「きれいな」画像を腫瘍があると誤分類することは、医師がその分類を確認するのにコストがかかるだけである。しかしながら、「腫瘍画像」をきれいであると誤分類することは、医師が全てのデータに手作業で目を通すことができない大きなデータセットでは、致命的になりかねない。そのようなシナリオでは、種々の誤分類リスクを、行列Ｍ内に適切に符号化することができる。能動トレーニングに関する大部分の研究と同様に、本発明の評価は、分類精度に基づく。したがって、本発明者らは、$M_{ij}=1$（ただし、$i \neq j$）のような、等しい誤分類コストを使用する。

【００４５】

サンプル選択

問い合わせ画像が与えられたとすると、サンプル選択手順は、ユーザーが与えなければならない応答の数を最小限に抑えるように、サンプル画像を選択する。本発明の枠組みでは、サンプル画像は既知のクラスに属する。その際、サンプル画像を選択するという問題は、サンプル画像として代表的な画像を選択することができる、「その問い合わせ画像のための有望なクラスを見つける」という問題と同じである。

【００４６】

問い合わせ画像及びサンプル画像を提示されるときに、ユーザーからの「イエス照合」応答は、実際に、問い合わせ画像そのもののクラスラベルを与える。

【００４７】

「ノー照合」応答は、ほとんど情報を与えない。データセットが１００クラスからなると仮定する。或る特定の問い合わせ－サンプル画像対に対してユーザーから「ノー照合」応答があると、問い合わせ画像が属することができる９９個の潜在的なクラスが依然として残される。この了解に基づいて、サンプル画像を選択する目的は、ユーザーからの「イエス照合」応答の尤度を最大することである。

【００４８】

サンプル画像（クラス）を選択することは、再び、選択された問い合わせ画像のための推定されたクラスメンバーシップ確率を用いることによって成し遂げることができる。表記を簡単にするために、問い合わせ画像分布｛$p_1, \ldots, p_k$）｝は、ソートされた順序において$p_1 \geqq p_2 \geqq \ldots \geqq p_k$のようになると仮定する。

【００４９】

その過程は、以下のように進む。クラス１から代表的なサンプル画像を選択し、ユーザー応答を得る。クラスｉ－１に対して「ノー照合」応答が得られる間、クラスｉからサンプル画像を選択して、ユーザーに提示する。これが、「イエス照合」応答が得られるまで続けられる。要求されるユーザー応答の数を最小限に抑えるために、そのような方式を通して、有望である可能性が高いクラスからのサンプル画像ほど、その過程において早期に選択される。

【００５０】

ラベル付けコスト

２値フィードバック環境において、ラベル付けのための各２値比較が一定のコスト（時間）を必要とすると仮定するのが妥当である。したがって、問い合わせ画像毎に、クラスラベルを得るためにかかるコストは、必要とされる２値比較の数に等しい。この数は、未知であるので、本発明では、代わりに、推定されるクラスメンバーシップ分布に基づいて、その期待値を求める。

【００５１】

その分布が、上記のようなソートされた順序において存在すると仮定される場合には、「イエス照合」応答を得るために予想されるユーザー応答の数は、次式（６）のようになる。

【００５２】

【数８】

$$C(x) = p_1^x + \sum_{j=2}^{k} (1 - p_1^x) \ldots (1 - p_{j-1}^x) \cdot p_j^x \cdot j \qquad (6)$$

【００５３】

これは、ユーザーラベル付けコストでもある。

【００５４】

本発明では、ラベル付けコストと同じ単位である、真のリスクを求めるのにかかる現実世界コストで、（Ｍをスケーリングすることによって）誤分類リスクをスケーリングすることができる。ここで、本発明では、真のリスクをアクティブセット内の誤分類の予想される数として選択し、アクティブセットサイズでＭをスケーリングすることによって、リスクを求める。Ｃ（ｘ）の選択と共に、これは結局、ユーザーからの各２値入力のコストをあらゆる誤分類と同等と見なすことになり、すなわち、本発明では、１つのラベル無し画像を正しく分類するために、ユーザーから１つの２値入力を得る。

【００５５】

終了判定基準

上記のＥＶＩに基づく目的関数は、興味を引くような終了判定基準に繋がる。本発明では、任意のラベル無し画像のための最大ＥＶＩが負であるときに、すなわち、次式が成立するときに、終了することができる。

【００５６】

【数９】

$$\mathrm{argmax}_{x \in A} V(x) < 0$$

【００５７】

本発明によるリスク及びコストの定義によれば、ＥＶＩの負の値は、ユーザーから入力される１つの２値が誤分類の数を１つも減らさないと予想されること、それゆえ、問い合わせが情報を得るだけの価値がないことを指示する。

【００５８】

特定の分野の知識が入手可能である場合には、代わりに、現実世界リスク及びラベル付けコストの異なる概念を利用できることに留意されたい。選択及び中止判定基準は、使用される特定の量をそのまま取り込む。

【００５９】

別の終了判定基準は、イエス及びノー照合応答に加えて、ユーザーから「無視」応答を得ることである。これにより、ユーザーは、１組のトレーニング画像から現在の問い合わせ画像を削除できるようになる（１４３）。そのような削除される画像は、雑音のある画像、複数の物体を示す画像、不鮮明な画像、及び分類器に学習してもらいたい任意の意味クラスに属さない画像の場合がある。

【００６０】
新たなクラスを開始する
多くの能動トレーニング法は、最初のトレーニングセットが全てのクラスからの例を含むことを制限的に仮定する。この仮定は、大多数の現実の問題にとって現実的でない。なぜなら、ユーザーが全てのクラスを有するトレーニングセットを明確に構成しなければならないので、管理を軽減するという本発明の目的を無にするためである。また、或る方法が長期にわたって機能するように期待される場合には、新たなクラスを取り扱うことは、不可欠である。したがって、本発明では、小さなシードセットから開始し、新たなクラスを動的に追加できるようにする。上記のサンプル選択法では、サンプル画像を示すことによって、「イエス照合」応答が得られるまでユーザーは問い合わせられる。しかしながら、問い合わせ画像が現在のトレーニングセット内に存在しないクラスに属する場合には、新たなクラスを開始するのに、数多くの問い合わせが必要とされることになる。
【００６１】
代わりに、本発明では、或る決まった少数、たとえば、５個の「ノー照合」応答が得られるときに、新たなクラスを開始する。良好なクラスモデルによって、予想される分布が、ラベル無し画像のクラスを正確にモデル化する。それゆえ、いくつかの最も有望なクラスに対して複数の「ノー照合」応答があることは、多くの場合に、以前に見かけていないクラスが存在することを指示する。
【００６２】
しかしながら、ラベル無し画像がトレーニングセット内に存在するクラスに属することも起こり得る。そのような場合には、新たなクラスを生成し、そのクラスをラベル無し画像に割り当てる結果として、過度のクラスタリングが生じる。これは、ユーザー入力と共に、ミニマックスカット手順後に、凝集クラスタリングによって、トレーニングされたクラスをマージすること（クラスタマージング）によって対処される。クラス間類似度スコアが凝集クラスタリング内の所定のしきい値よりも大きい場合には、そのクラスタリングは、２つのクラスを結合する。この場合、２つのクラスからの２つのトレーニング画像がユーザーに対して表示され、２つのクラスがマージされるべきであるか否かを尋ねるときに、その応答がイエス照合である場合には、２つのクラスターはマージされる。
【００６３】
凝集クラスタリングにおける基本的な発想は、最も高い類似度（連関値）１（C_i，C_j）を有する２つのクラスターを繰り返しマージすることである。ミニマックスクラスタリングの場合、連関関数は、次式のようになる。
【００６４】
【数１０】

$$1(C_i, C_j) = s(C_i, C_j) / (s(C_i, C_i)s(C_j, C_j))$$

【００６５】
ただし、ｓはクラスター類似度スコア（次式）を示す。
【００６６】
【数１１】

$$s(C_i, C_j) = \sum_{x \in C_i} \sum_{y \in C_j} K(x, y)$$

【００６７】
ここで、Ｋは２つの画像ｘとｙとの間の類似度を取り込むカーネル関数である（ＳＶＭを用いる分類の場合にも、同じカーネル関数が用いられる）。
【００６８】
本発明の方法では、ユーザーフィードバックを繰り返す度に、クラスター連関値を評価する。クラスターC_i及びC_jの場合に、最大連関値（クラスターが重なることを指示す

る）であり、０．５のしきい値より大きい場合には、クラスターＣ$_i$及びＣ$_j$からの２つの画像を示すことによって、ユーザーに問い合わせる。

【００６９】

「イエス照合」応答の場合、結果として２つのクラスターがマージされる。本発明では、ユーザーフィードバックを有するので、本発明の方法は、教師なしクラスタリングよりもはるかに簡単である。したがって、この方法は、用いられる特定のしきい値の影響を比較的受けにくく、受ける雑音が少ない。また、本発明では、繰り返す度に連関値を求める必要はないことにも留意されたい。簡単な増分計算しか必要とされない。要するに、新たなクラスが迅速に開始され、誤ったクラスは、ほとんどユーザーフィードバックを用いることなくクラスタマージングによって訂正される。

【００７０】

計算に関する検討

サイズＮ及びｋクラスのアクティブセットの場合に、図２の本発明の手順における各問い合わせ繰り返しの計算の複雑さは、Ｏ（Ｎ^{2}ｋ^3）である。それは、小さな問題の場合に十分に機能するが、規模が大きくなるほど、そのコストが実用的でなくなる可能性がある。それゆえ、本発明ではいくつかの近似を用いて、計算費用を著しく低減し、数多くのクラスを有する大きな問題の場合でも、その実施を効率的にする。

【００７１】

期待値計算

上記の手順では、予想されるリスクを推定することは、費用がかかる。本発明では、ラベル無し画像毎に、あり得るクラスのいずれかに画像が属することができると仮定して、分類器をトレーニングする必要がある（図２の手順の４行目）。これは、数多くのクラスが存在するときに時間がかかる可能性がある。

【００７２】

これを克服するために、本発明では、以下の観測を行なう。ラベル無し画像の推定確率分布が与えられたとすると、低い確率値を割り当てられるクラスに属する可能性は低く、すなわち、その画像は、最も高い推定確率を有するクラスに属する可能性が最も高い。

【００７３】

したがって、全ての取り得るクラスを一巡する代わりに、本発明者らは、最も可能性の高いクラスだけを一巡することができる。詳細には、上から２つの最も可能性の高いクラスは、判別情報の大部分を含み、一方、確率値が小さくなるほど、ほとんど情報を含まないので、本発明では、上から２つの最も可能性の高いクラスだけを一巡する。

【００７４】

そのような近似は、或る程度まで、推定されるモデルが正確であることを当てにしており、それは、計算しやすくするために多くの場合に行なわれる楽観的な仮定を含む。さらに、本発明では、式（１）に対する近似として、ラベル無し画像に関する予想されるリスクを計算するために（９行目）、同じ「上から２つ」の近似を用いることができる。

【００７５】

リスクを推定するためのクラスタリング

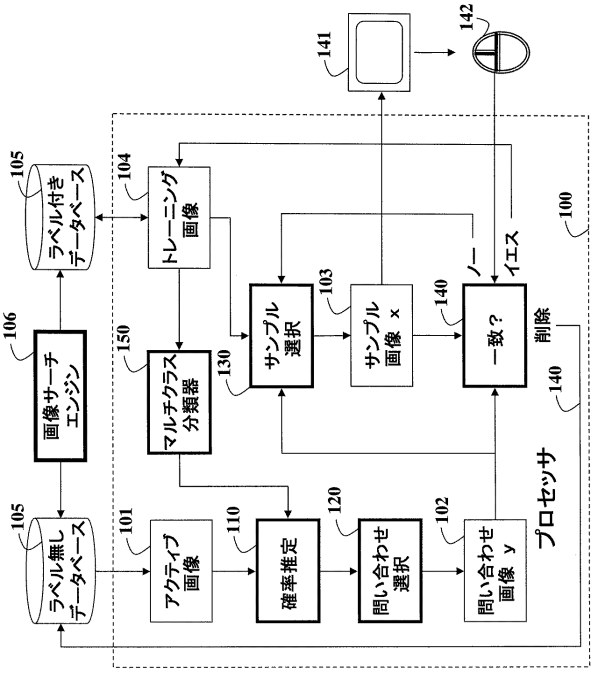
上記の手順では、完全なアクティブセットにおいて、リスクを推定する必要がある。代わりに、本発明では、最初に、カーネルｋ平均手順を用いて、アクティブセット内のラベル無し画像をクラスタリングする。その後、本発明では、各クラスターから１つの代表的な（図心に最も近い）画像を選択することによって、新たなラベル無し画像セットを形成し、この削減されたセットに関してリスクを推定する。

【００７６】

最初に、一度だけクラスタリングを実行し、問い合わせが繰り返される度に実行しない。本発明の実施態様では、クラスターの数をアクティブセットサイズの１／１００に固定する。この近似は、能動的に選択された画像を稀に（時間のうちの５％未満）しか変更せず、推定されるリスク値、及び将来の分類精度において無視できる程度の差しかもたらさない。

【００７７】
用いられる別の近似は、アクティブセットをサンプリングして、ＥＶＩを実行するために得られるセットを小さくすることである。不確定サンプリングのような、効率的な能動選択ヒューリスティックを利用して、小さなセットを形成することができる。不確定サンプリング手順を用いるとき、本発明では、最初に、アクティブセットから約５０画像の小さなセットをサンプリングし、その後、ＥＶＩを用いて、この小さくなったセットから画像を選択する。
【００７８】
上記の近似によれば、各問い合わせ繰り返しの複雑さは、Ｏ（Ｎｋ^{2}）であり、元のバージョンよりも大きく改善される。これは、従来技術における能動的選択において多くの場合に観測される３乗スケーリングよりもはるかに良好である。
【００７９】
本発明では、ユーザーからの２値フィードバックのみを必要とする新たなマルチクラス能動トレーニングの枠組みを記述する。本発明の方法は、ユーザートレーニング時間及び労力を大きく削減する。その方法は、母集団不均衡及び雑音のような現実世界の問題に対してロバストである。
【００８０】
図４は、従来のインタラクティブ分類器トレーニングを示しており、ユーザーは、数百のカテゴリを含む可能性がある大きなデータセットから、問い合わせ画像のための１つのクラス３０１を入力する必要がある。対照的に、図３に示されるように、ユーザーは、問い合わせ画像及びサンプル画像が一致するか否か（イエス／ノー３０２）、それゆえ、同じクラスのメンバーであるか否かを指示するだけである。
【００８１】
本発明を、好ましい実施の形態の例として説明してきたが、本発明の精神及び範囲内で他のさまざまな適合及び変更を行えることが理解されるべきである。したがって、本発明の真の精神及び範囲内に入るすべての変形及び変更を包含することが、添付の特許請求の範囲の目的である。

【図１】

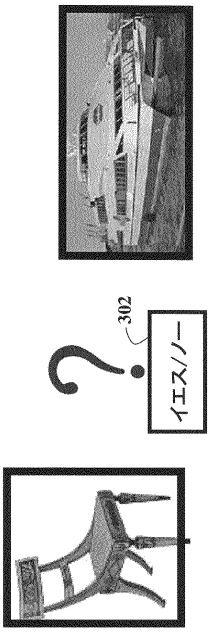


【図２】

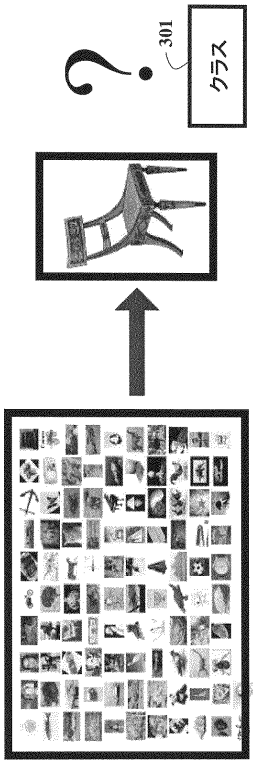
Input: ラベル付きセットL, アクティブプールA, コスト行列M

$\mathcal{L}^0 := \mathcal{L}; \mathcal{A}^0 := \mathcal{A}$
for round $r = 0$ **to** $n - 1$ **do**
  foreach image $x_i \in \mathcal{A}^{(r)}$ **do**
    for class $y_l = 1$ **to** k **do**
      $\mathrm{L}^{(r)} \cup \{x_i, y_l\}$を用いてマルチクラス分類器をトレーニングする
      アクティブプール$\mathrm{A}^{(r)}$内の画像のためのクラスメンバーシップ確率を推定する
      アクティブプール$R^{(x_i,y_l)}$においてリスクを計算する
    end
    予想されるリスクを計算する $(\mathcal{L}^{r'} = \mathcal{L}^r \cup \{x_i\})$
    $\mathcal{R}_{\mathcal{L}^{r'}} = \sum_l P(y_l = l) \cdot R^{(x_i,l)}$
    予想される注釈コストC(x_i)を計算する
  end
  Find image $x^* = \mathrm{argmin}_{x_i \in \mathcal{A}^{(r)}} \mathcal{R}_{\mathcal{L}^{r'}} + \mathcal{C}(x_i)$
  Find $V(x^*)$ using Eqn. (4)
  if $V(x^*) > 0$ **then**
    真のラベルk^*が得られるまで、問い合わせ画像x^*及び有望なサンプル画像をユーザーに問い合わせる
    Set $\mathcal{L}^{(r+1)} := \mathcal{L}^{(r)} \cup \{x^*, k^*\}$; and $\mathcal{A}^{(r+1)} := \mathcal{A}^{(r)} \setminus \{x^*\}$
  else return $\mathcal{L}^{(n)} = \mathcal{L}^{(r)}$
end

【図３】



【図４】



フロントページの続き

(74)代理人 100122437
弁理士 大宅 一宏
(74)代理人 100147566
弁理士 上田 俊一
(74)代理人 100161171
弁理士 吉田 潤一郎
(74)代理人 100161115
弁理士 飯野 智史
(72)発明者 ファティー・ポリクリ
アメリカ合衆国、マサチューセッツ州、ウォータータウン、マウント・オーバーン・ストリート７１２
(72)発明者 アジェイ・ジョシ
アメリカ合衆国、ミネソタ州、ミネアポリス、セカンド・アベニュー・エスイー ５２０、アパートメント ２０５

審査官 ▲広▼島 明芳

(56)参考文献 特開２００６－２４４４６２（ＪＰ，Ａ）
米国特許出願公開第２００８／０１１８１５１（ＵＳ，Ａ１）
特開２００１－１２８０００（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｔ ７／００
Ｇ０６Ｎ ３／００