図 9.16 「金閣寺」画像.

図 9.17 「ラーメン」画像.

図 9.18 「入学式」画像.

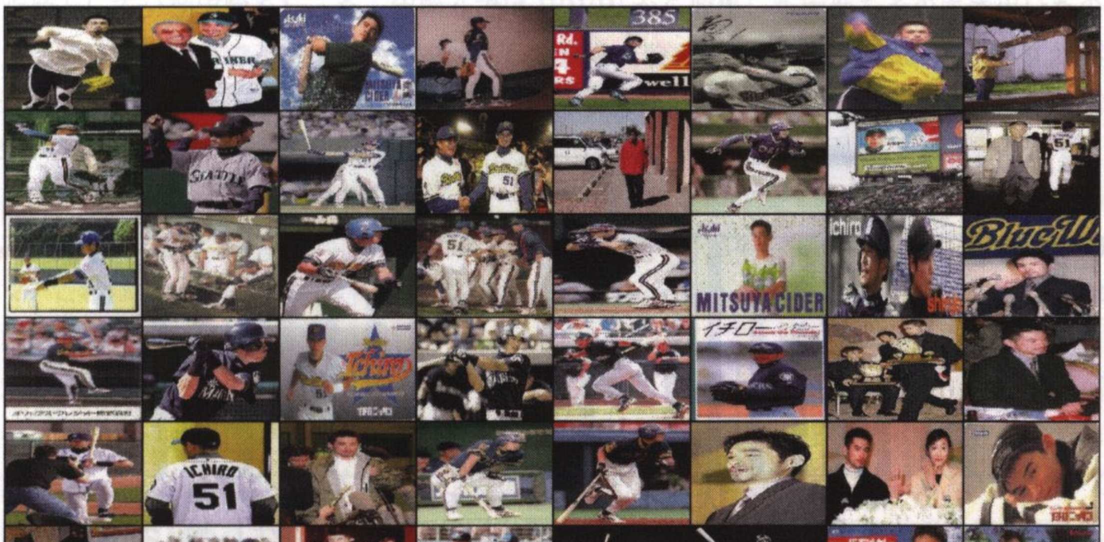

図 9.19 「イチロー」画像.

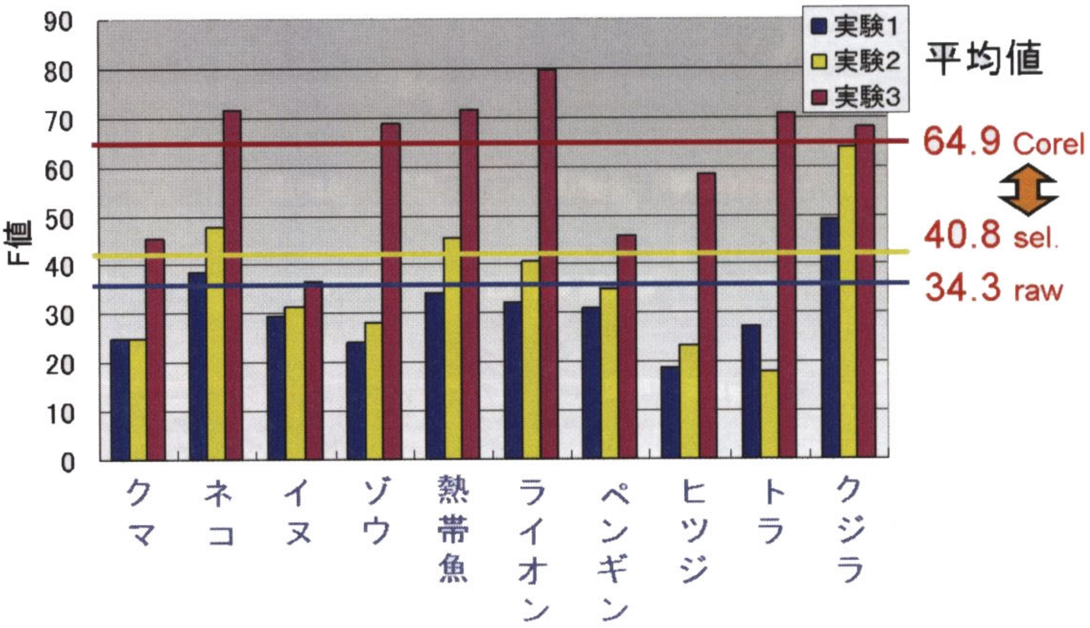

図 9.20 実験 1, 2, 3 の結果 (方法 2).

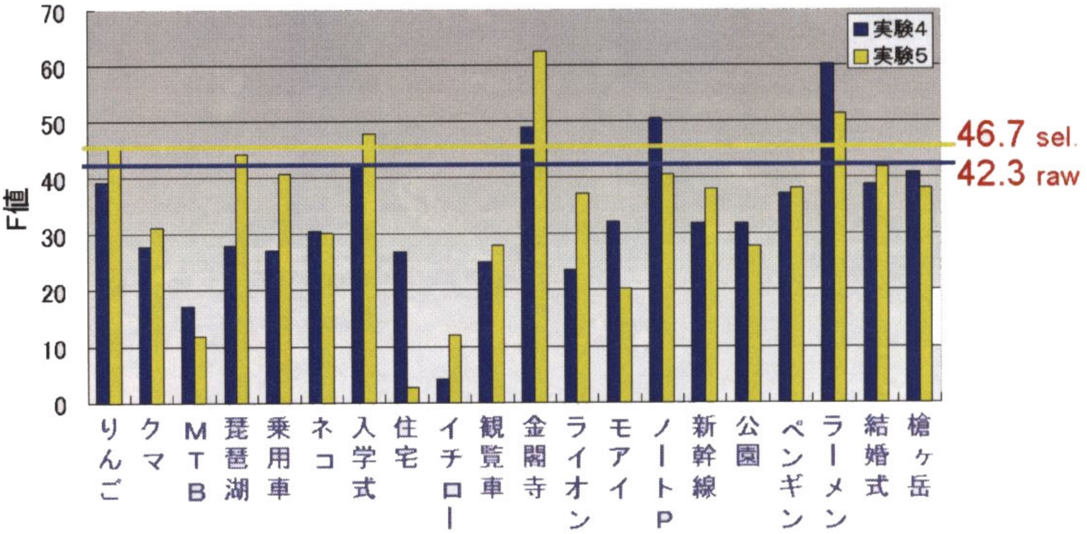

図 9.21 実験 4, 5 の結果 (方法 2).

9.4.2 一般画像の分類

次に，WWWから収集した画像を学習画像として,「一般画像」の分類実験を行った．最初に「一般画像」のテスト画像としてCorel画像を用いて，WWW収集画像を学習画像として分類実験を行った．しかし，前節の実験で見たように，Corel画像は互いに類似した画像が多く，多様である実世界の「一般画像」の典型的な画像であるとは言い切れない．そこで，さらに，WWW収集画像およびCorel画像から人手で各クラスの典型的な，そして，なるべく多様な画像を50枚ずつ選び出し，それをテスト画像として分類を行った．一般画像の分類実験として行った実験の一覧を表9.8に示す．

表 9.8 分類実験.

実験		学習画像		テスト画像	
no.	クラス数	枚数	適合率 (%)	方法	入手元
6	10	4548	68.2	別画像 (50枚)	Corel
7	10	3102	100†	別画像 (50枚)	Corel
8	20	5694	61.2	別画像 (50枚)	WWW, 一部Corel
9	20	3485	100†	別画像 (50枚)	WWW, 一部Corel

†人手で適合画像のみを選択.

実験6, 実験7では，実験1, 実験2でWWWから収集した画像を学習画像とし，Corel画像をテスト画像として分類実験を行った．実験6ではWWWから収集した画像をそのまま用いたのに対して，実験7では人手で不適切な画像を除去してから実験を行った．結果を表9.9，図9.22および表9.10，図9.23にそれぞれ示す．また，図9.26に平均的に最も結果のよかった方法2による実験6, 7の2つの実験の結果をまとめて示した．ランダムの場合にF値が10.0になることを考えると，実験6の方法3による結果のF値24.1，実験8の同じく方法3による27.6はそれよりはかなり良い結果であった．ただし，図9.26に示したように，値自体は実験1，実験2に比べてそれぞれ10程度低く，これはCorel画像がほとんど野生の動物であるのに対して，WWWから収集した画像には動物園の画像が多いという性質の違いによるところが大きいと考えられる．

個々のクラスの分類結果に関しては,「イヌ」「熱帯魚」「ライオン」「ペンギン」「クジラ」が比較的良い分類結果を残していた一方，それ以外はランダムの場合と同程度であった．その中でも「クジラ」に関しては，どちらにもホエールウォッチングの船上から撮影した写真がWWW画像，Corel画像の両方にある程度含まれていたので，平均値に比べると良い値である50前後の値が得られた．

実験8，実験9では，一般の実世界画像に対する分類能力の評価として，WWW収集画像とCorel画像から人手で各クラスの典型的な，そして，なるべく多様な画像を50枚ずつ選び出し，それをテスト画像として分類を行った．本来なら標準的な評価用データセットが存在していれば，それを用いて分類性能の評価を行うべきであるが，一般に実世界画像に関してはそのような評価用データというものは存在しない．そこで，クラスによって異なるが，最大Corel画像から15枚，残りを

WWW 収集画像から人手によって取出して，各クラスについて 50 枚の評価用画像セットを構築した．なお，付録 Bの図 B.1〜B.20に使用した 1000 枚の評価画像すべてを示してある．付録 B中においては，実験 8 において正しく分類された画像を赤い枠で囲ってあり，誤って分類された画像については，分類されたクラスを各画像の下に引いた下線の色の組合わせで表現している．

実験 8 では，実験 4 で用いた WWW から収集しただけの画像データを用いて，評価用データに対する分類を行った．実験 9 では，実験 5 で用いた人手によって不適合画像を除去した後の画像データを用いて，分類実験を行った．結果をそれぞれ，表 9.11，図 9.24および表 9.12，図 9.25に示す．また，図 9.27に方法 2 による実験 8, 9 の結果を示した．赤い物体が画像の大部分を占めている「りんご」，湖面と空からなる「琵琶湖」，黒い制服を着た人がたくさん写っている「入学式」，池と金色の建物からなる「金閣寺」，灰色の線路に白い車体の「新幹線」，緑の芝生の「公園」，青い空と灰色の岩山の「槍ヶ岳」などの傾向のはっきりとしている画像は，うまく分類できていることが分かる．

実験 8 では方法 2 で，再現率 39.8%，適合率 50.7%で 44.6 の F 値が得られた．これは，一般画像に対する認識結果としては極めて良い結果である．我々の目的とする自動分類は，学習画像の収集に人手を一切介入させないで，高い分類精度を得ることであるので，まさにこの実験 8 は我々の目的としている分類実験である．それで，これだけの結果と得られたということは類似画像検索手法による画像分類が一定の有効性を持つ方法であることを示していると言える．

実験 9 は，実験 5 で用いた適合画像のみで分類実験を行った．その結果，方法 2 による場合は F 値が 47.0 であり，3 程度しか向上しなかった．このことは，今回実験した方法においては，学習画像集合にノイズ画像が含まれていても，それほど大きく認識結果に影響を与えないということを示している．

同じ 20 クラスで行った実験 4，実験 5 は実験 8，実験 9 よりも 20 程度低い F 値で，人手でテスト画像を「典型的な」画像に限定したことによってその分結果が良くなったということである．究極的にはどんな画像でも分類できることが望ましいが，現在においてはこのような「典型的な」画像の分類でも 100%近くにすることは容易ではない．当面は，このような「典型的な」画像の分類精度を上げることに重要である．

実験 8, 実験 9 の分類結果については，どのクラスの画像が何枚どのクラスに分類されたかを示す confusion matrix を示す．実験 8 方法 2，実験 9 方法 2，実験 8 方法 3，実験 9 方法 3 による分類結果の confusion matrix を表 9.13，表 9.14，表 9.15，表 9.16にそれぞれ示す．表の縦方向はテスト画像のクラス名を示し，横方向は分類先のクラスを示す．つまり，表 9.13では，「りんご」画像 50 枚のうち，26 枚が正しく「りんご」クラスに分類され，それ以外は「公園」クラスに 7 枚，「結婚式」クラスに 6 枚，その他のクラスに 11 枚が分類されたということを示している．この confusion matrix からは，どのクラスの画像がどのクラスに間違って分類されやすいかを知ることが出来る．表 9.13からは，例えば，18 枚もの「クマ」画像が「ネコ」に，14 枚もの「乗用車」画像が「新幹線」に分類されていることが分かる．前者はどちらも黒い物体，後者はどちらも白い物体であることが原因になっていると考えられる．

方法 2 による結果では，実験 8 でも実験 9 でも，分類された画像枚数がクラスによって大きくば

らつきがあることが分かる.「イチロー」にはどちらも 0 枚である一方,「公園」には実験 8 で 105 枚，実験 9 で 216 枚も画像が分類されていることが分かる．これは,「公園」には空や木や緑の芝生など，他のクラスにも存在しそうな物体が多く含まれているのに対して,「イチロー」が画像自体にこれといった典型パターンがなく，様々な色が含まれている．そのため方法 2 での色の特徴量が広範囲に分布してしまって，25 ブロックの最小距離の合計が「イチロー」クラスに対して最小になるような画像が存在しなかったことが考えられる.

方法 3 に対する結果では，方法 2 の時ほどはクラス間によって分類画像枚数の差が大きくはない. ただし,「イチロー」に関しては，実験 8，実験 9 共に分類枚数は少ない．ただし，それぞれの実験で 6 枚，7 枚と方法 2 に比べると多くなっている．これは方法 3 では，色に加えてテクスチャや領域の形状も特徴量として考慮されていることが原因として考えられる．全体的な傾向としては，方法 2 の場合ほど大きな偏りはないが，誤分類のパターンは類似している傾向がうかがえる．また，方法 3 と方法 4 についても，誤分類の傾向はよく類似しているが，全体的には方法 4 の様が低い結果に終わっている.

表 9.9 実験 6. 10 クラス (不適合画像含む) の WWW 収集画像を用いて Corel 画像を分類した結果.

	方法 1			方法 2			方法 3			方法 4		
クラス名	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値
クマ	4.0	5.9	4.8	8.0	15.4	10.5	8.0	16.0	10.7	8.0	13.3	10.0
ネコ	6.1	14.3	8.6	4.1	33.3	7.3	8.2	10.5	9.2	2.0	3.8	2.7
イヌ	22.0	22.9	22.4	24.0	14.8	18.3	26.0	15.5	19.4	26.0	23.2	24.5
ゾウ	10.0	11.4	10.6	34.0	34.7	34.3	12.0	13.6	12.8	14.0	13.5	13.7
熱帯魚	12.0	21.4	15.4	22.0	61.1	32.4	24.0	31.6	27.3	4.0	14.3	6.2
ライオン	8.0	7.3	7.6	30.0	19.5	23.6	22.0	22.9	22.4	12.0	14.3	13.0
ペンギン	12.0	15.4	13.5	26.0	19.7	22.4	30.0	30.0	30.0	18.0	14.3	15.9
ヒツジ	2.0	3.7	2.6	8.0	23.5	11.9	6.0	9.4	7.3	8.0	6.6	7.2
トラ	6.0	13.0	8.2	4.0	7.4	5.2	12.0	16.2	13.8	20.0	18.2	19.0
クジラ	80.0	22.2	34.8	86.0	32.6	47.3	84.0	40.8	54.9	64.0	32.0	42.7
平均値	16.2	13.8	**14.9**	24.6	26.2	**25.4**	23.2	20.7	**21.9**	17.6	15.3	**16.4**

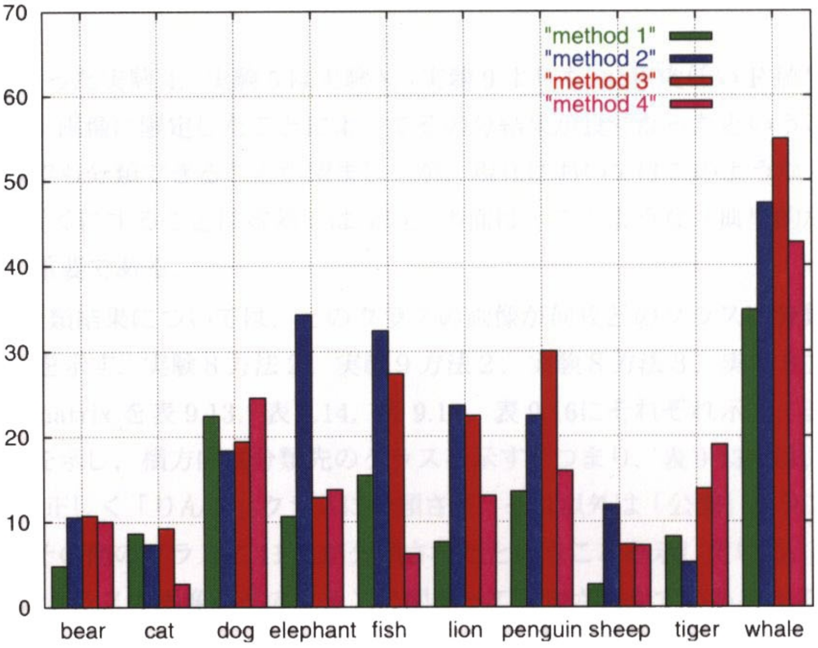

図 9.22 実験 6. 10 クラス (不適合画像含む) の WWW 収集画像を用いて Corel 画像を分類した結果.

表 9.10 実験 7. 10 クラス (不適合画像あり) の WWW 収集画像を用いて Corel 画像を分類した結果.

	方法 1			方法 2			方法 3			方法 4		
クラス名	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値
クマ	6.0	20.0	9.2	4.0	40.0	7.3	8.0	26.7	12.3	6.0	10.0	7.5
ネコ	4.1	9.1	5.6	6.1	42.9	10.7	6.1	9.1	7.3	6.1	7.9	6.9
イヌ	38.0	18.1	24.5	58.0	21.3	31.2	38.0	19.0	25.3	28.0	18.4	22.2
ゾウ	12.0	35.3	17.9	16.0	25.8	19.8	14.0	17.9	15.7	10.0	12.2	11.0
熱帯魚	30.0	40.5	34.5	30.0	46.9	36.6	42.0	41.2	41.6	10.0	20.8	13.5
ライオン	18.0	21.4	19.6	36.0	27.3	31.0	24.0	22.2	23.1	12.0	16.7	14.0
ペンギン	40.0	30.8	34.8	48.0	25.5	33.3	30.0	27.3	28.6	22.0	13.8	16.9
ヒツジ	0.0	0.0	0.0	4.0	18.2	6.6	6.0	11.5	7.9	8.0	8.5	8.2
トラ	0.0	0.0	0.0	10.0	45.5	16.4	4.0	13.3	6.2	16.0	33.3	21.6
クジラ	90.0	24.3	38.3	86.0	40.6	55.1	88.0	39.6	54.7	66.0	32.0	43.1
平均値	23.8	20.0	**21.7**	29.8	33.4	**31.5**	26.0	22.8	**24.3**	18.4	17.4	**17.9**

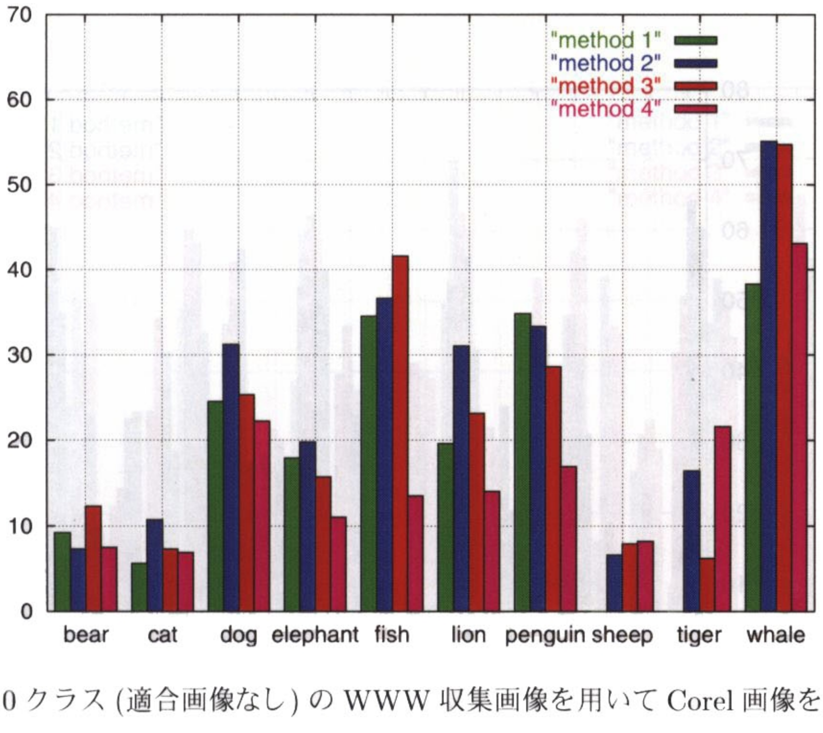

図 9.23 実験 7. 10 クラス (適合画像なし) の WWW 収集画像を用いて Corel 画像を分類した結果.

表 9.11 実験 8. 20 クラス (不適合画像含む) の WWW 画像による典型画像の分類結果.

		方法 1			方法 2			方法 3			方法 4		
no.	クラス名	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値	再現率	適合率	F 値
1	りんご	46.0	88.5	60.5	52.0	83.9	64.2	42.0	63.6	50.6	40.0	45.5	42.6
2	クマ	4.0	4.5	4.3	14.0	17.9	15.7	26.0	20.3	22.8	18.0	14.1	15.8
3	マウンテンバイク	0.0	0.0	0.0	8.0	40.0	13.3	14.0	28.0	18.7	10.0	16.1	12.3
4	琵琶湖	48.0	23.1	31.2	64.0	31.1	41.8	56.0	41.8	47.9	46.0	27.4	34.3
5	乗用車	2.0	100.0	3.9	16.0	88.9	27.1	38.0	54.3	44.7	24.0	50.0	32.4
6	ネコ	24.0	17.6	20.3	54.0	26.0	35.1	40.0	23.5	29.6	42.0	25.9	32.1
7	入学式	52.0	37.1	43.3	68.0	47.9	56.2	78.0	48.1	59.5	70.0	42.7	53.0
8	住宅	0.0	0.0	0.0	4.0	50.0	7.4	22.0	55.0	31.4	28.0	77.8	41.2
9	イチロー	0.0	0.0	0.0	0.0	0.0	0.0	6.0	50.0	10.7	4.0	22.2	6.8
10	観覧車	16.0	66.7	25.8	22.0	73.3	33.8	40.0	55.6	46.5	36.0	43.9	39.6
11	金閣寺	74.0	43.0	54.4	78.0	41.5	54.2	74.0	42.0	53.6	62.0	44.9	52.1
12	ライオン	4.0	66.7	7.5	18.0	100.0	30.5	10.0	33.3	15.4	14.0	25.0	17.9
13	モアイ	16.0	88.9	27.1	42.0	91.3	57.5	50.0	62.5	55.6	44.0	50.0	46.8
14	ノートパソコン	22.0	91.7	35.5	30.0	93.8	45.5	46.0	79.3	58.2	50.0	75.8	60.2
15	新幹線	42.0	21.9	28.8	66.0	31.7	42.9	60.0	39.5	47.6	40.0	30.3	34.5
16	公園	66.0	23.2	34.4	52.0	24.8	33.5	28.0	20.6	23.7	24.0	19.0	21.2
17	ペンギン	20.0	16.7	18.2	32.0	37.2	34.4	18.0	18.8	18.4	22.0	26.2	23.9
18	ラーメン	60.0	40.5	48.4	60.0	61.2	60.6	56.0	50.0	52.8	46.0	41.1	43.4
19	結婚式	42.0	31.3	35.9	46.0	45.1	45.5	32.0	34.8	33.3	36.0	40.0	37.9
20	槍ヶ岳	70.0	28.2	40.2	70.0	29.2	41.2	66.0	40.2	50.0	62.0	40.8	49.2
	平均値	30.4	39.5	**34.4**	39.8	50.7	**44.6**	40.1	43.1	**41.5**	35.9	37.9	**36.9**

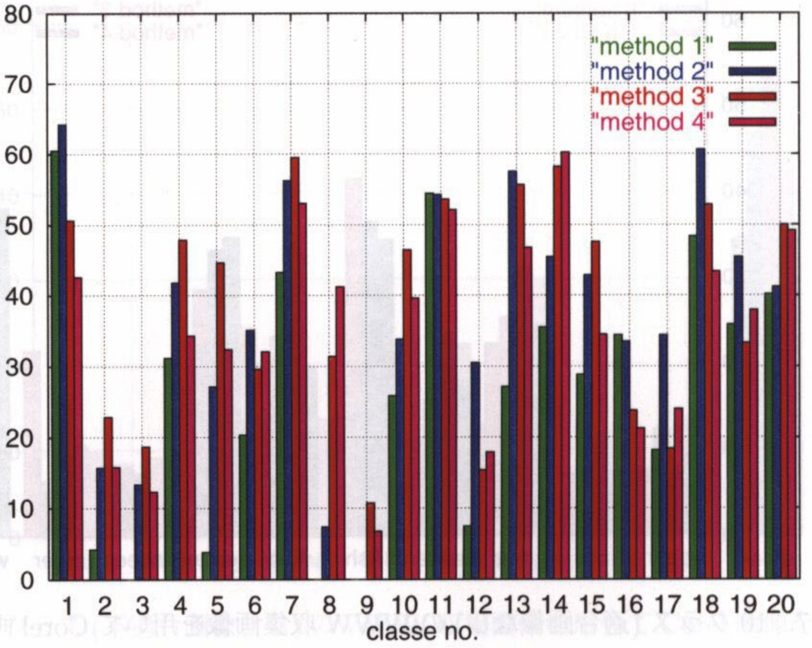

図 9.24 実験 8. 20 クラス (不適合画像含む) の WWW 画像による典型画像の分類結果.

表 9.12 実験 9. 20 クラス (不適合画像なし) の WWW 画像による典型画像の分類結果.

no.	クラス名	方法 1 再現率	方法 1 適合率	方法 1 F 値	方法 2 再現率	方法 2 適合率	方法 2 F 値	方法 3 再現率	方法 3 適合率	方法 3 F 値	方法 4 再現率	方法 4 適合率	方法 4 F 値
1	りんご	22.0	84.6	34.9	36.0	90.0	51.4	36.0	58.1	44.4	44.0	56.4	49.4
2	クマ	0.0	0.0	0.0	8.0	66.7	14.3	18.0	26.5	21.4	14.0	16.7	15.2
3	マウンテンバイク	2.0	100.0	3.9	4.0	66.7	7.5	10.0	29.4	14.9	10.0	22.7	13.9
4	琵琶湖	48.0	50.0	49.0	52.0	47.3	49.5	66.0	52.4	58.4	56.0	35.0	43.1
5	乗用車	6.0	100.0	11.3	26.0	81.2	39.4	36.0	46.2	40.4	36.0	48.6	41.4
6	ネコ	34.0	14.3	20.1	62.0	18.3	28.3	50.0	23.1	31.6	44.0	20.8	28.2
7	入学式	58.0	28.7	38.4	76.0	48.7	59.4	88.0	47.8	62.0	78.0	47.0	58.6
8	住宅	0.0	0.0	0.0	4.0	100.0	7.7	30.0	65.2	41.1	26.0	65.0	37.1
9	イチロー	0.0	0.0	0.0	0.0	0.0	0.0	8.0	57.1	14.0	6.0	33.3	10.2
10	観覧車	6.0	100.0	11.3	16.0	100.0	27.6	26.0	86.7	40.0	26.0	76.5	38.8
11	金閣寺	40.0	64.5	49.4	60.0	83.3	69.8	64.0	60.4	62.1	60.0	60.0	60.0
12	ライオン	4.0	100.0	7.7	10.0	100.0	18.2	10.0	38.5	15.9	12.0	46.2	19.0
13	モアイ	16.0	100.0	27.6	26.0	86.7	40.0	44.0	66.7	53.0	36.0	56.2	43.9
14	ノートパソコン	24.0	100.0	38.7	32.0	94.1	47.8	42.0	87.5	56.8	48.0	85.7	61.5
15	新幹線	16.0	15.7	15.8	58.0	49.2	53.2	54.0	40.3	46.2	52.0	40.0	45.2
16	公園	88.0	15.0	25.7	82.0	19.0	30.8	58.0	23.0	33.0	46.0	21.1	28.9
17	ペンギン	6.0	25.0	9.7	24.0	44.4	31.2	18.0	22.5	20.0	26.0	28.9	27.4
18	ラーメン	74.0	33.9	46.5	64.0	45.1	52.9	60.0	42.3	49.6	58.0	36.2	44.6
19	結婚式	50.0	30.5	37.9	52.0	36.6	43.0	40.0	31.2	35.1	36.0	40.9	38.3
20	槍ヶ岳	68.0	30.6	42.2	78.0	31.0	44.3	84.0	52.5	64.6	72.0	45.6	55.8
	平均値	28.1	49.6	**35.9**	38.5	60.4	**47.0**	42.1	47.9	**44.8**	39.3	44.1	**41.6**

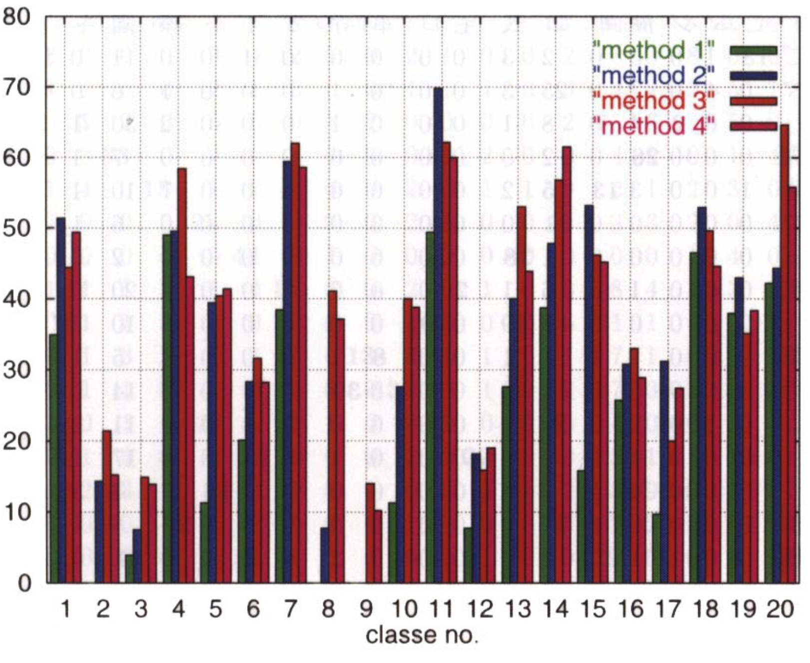

図 9.25 実験 9. 20 クラス (不適合画像なし) の WWW 画像による典型画像の分類結果.

表 9.13 実験 8：方法 2 による実験結果の confusion matrix

	りんご	クマ	マウンテンバイク	琵琶湖	乗用車	ネコ	入学式	住宅	イチロー	観覧車	金閣寺	ライオン	モアイ	ノート	新幹線	公園	ペンギン	ラーメン	結婚式	槍ヶ岳	**枚数**	**再現率**
りんご	**26**	3	1	0	0	0	2	0	0	0	0	0	0	0	1	7	1	3	6	0	50	52.0
クマ	0	**7**	0	2	0	18	1	0	0	0	4	0	0	0	7	4	3	0	1	3	50	14.0
マウンテンバイク	0	2	**4**	9	0	4	0	1	0	1	5	0	0	0	6	12	2	2	0	2	50	8.0
琵琶湖	0	0	0	**32**	0	1	0	0	0	1	0	0	0	0	0	3	2	0	0	11	50	64.0
乗用車	1	0	0	9	**8**	1	5	0	0	0	4	0	0	0	14	4	1	1	1	1	50	16.0
ネコ	1	6	0	2	0	**27**	0	0	0	0	2	0	0	0	2	1	1	5	1	2	50	54.0
入学式	0	1	0	0	0	2	**34**	0	0	0	1	0	0	0	1	1	0	3	5	2	50	68.0
住宅	0	0	1	7	0	3	0	**2**	0	0	9	0	1	0	8	6	3	1	1	8	50	4.0
イチロー	1	2	1	2	0	3	17	0	**0**	0	2	0	0	0	5	6	3	0	5	3	50	0.0
観覧車	1	0	0	7	0	4	3	0	0	**11**	2	0	0	0	4	2	2	0	2	12	50	22.0
金閣寺	0	1	1	0	0	1	0	0	0	0	**39**	0	0	0	1	7	0	0	0	0	50	78.0
ライオン	0	8	0	2	0	11	1	0	0	0	2	**9**	0	0	0	8	2	2	3	2	50	18.0
モアイ	0	1	0	8	0	1	0	0	0	1	2	0	**21**	0	0	6	0	2	0	8	50	42.0
ノートパソコン	0	3	0	2	1	2	0	1	0	0	1	0	0	**15**	6	1	5	0	0	13	50	30.0
新幹線	0	0	0	3	0	6	0	0	0	0	0	0	0	0	**33**	5	2	0	0	1	50	66.0
公園	0	1	2	4	0	0	0	0	0	0	9	0	1	0	1	**26**	0	0	0	6	50	52.0
ペンギン	0	2	0	4	0	8	2	0	0	0	0	0	0	0	5	3	**16**	0	0	10	50	32.0
ラーメン	1	2	0	2	0	5	1	0	0	0	1	0	0	0	2	2	0	**30**	3	1	50	60.0
結婚式	0	0	0	5	0	5	5	0	0	1	7	0	0	0	4	0	0	0	**23**	0	50	46.0
槍ヶ岳	0	0	0	3	0	2	0	0	0	0	4	0	0	1	4	1	0	0	0	**35**	50	70.0
分類枚数	31	39	10	103	9	104	71	4	0	15	94	9	23	16	104	105	43	49	51	120	1000	**39.8**
適合率	83	17	40	31	88	25	47	50	0	73	41	100	91	93	31	24	37	61	45	29	**50.7**	**44.6**

表 9.14 実験 9：方法 2 による実験結果の confusion matrix

	りんご	クマ	マウンテンバイク	琵琶湖	乗用車	ネコ	入学式	住宅	イチロー	観覧車	金閣寺	ライオン	モアイ	ノート	新幹線	公園	ペンギン	ラーメン	結婚式	槍ヶ岳	**枚数**	**再現率**
りんご	**18**	1	0	0	0	2	3	0	0	0	0	0	1	0	0	11	0	10	4	0	50	36.0
クマ	0	**4**	0	2	0	25	3	0	0	0	1	0	0	0	1	6	0	2	5	1	50	8.0
マウンテンバイク	0	0	**2**	4	0	8	1	0	0	0	1	0	0	0	2	30	1	1	0	0	50	4.0
琵琶湖	0	0	0	**26**	0	2	0	0	0	0	0	0	0	0	0	7	1	0	0	14	50	52.0
乗用車	1	0	0	3	**13**	5	2	0	0	0	0	0	0	0	7	10	1	2	1	5	50	26.0
ネコ	0	0	0	0	0	**31**	0	0	0	0	0	0	0	0	0	6	1	7	3	2	50	62.0
入学式	0	0	0	0	0	1	**38**	0	0	0	0	0	0	0	0	2	0	4	4	1	50	76.0
住宅	0	0	0	1	0	6	1	**2**	0	0	0	0	0	0	4	20	0	0	4	12	50	4.0
イチロー	0	0	0	0	0	3	19	0	**0**	0	0	0	0	0	3	10	3	2	8	2	50	0.0
観覧車	0	0	0	7	1	7	1	0	0	**8**	1	0	0	0	1	5	1	2	6	10	50	16.0
金閣寺	0	0	0	0	0	1	0	0	0	0	**30**	0	0	0	0	14	0	2	1	2	50	60.0
ライオン	0	1	0	2	0	20	1	0	0	0	0	**5**	0	0	0	11	2	4	3	1	50	10.0
モアイ	0	0	0	1	0	1	0	0	0	0	1	0	**13**	0	0	17	1	3	2	11	50	26.0
ノートパソコン	0	0	0	4	1	2	0	0	0	0	0	0	0	**16**	7	4	3	0	0	13	50	32.0
新幹線	0	0	0	0	1	11	0	0	0	0	0	0	0	0	**29**	5	1	0	2	1	50	58.0
公園	0	0	1	1	0	2	0	0	0	0	2	0	1	0	0	**41**	0	0	0	2	50	82.0
ペンギン	0	0	0	3	0	17	3	0	0	0	0	0	0	0	2	3	**12**	0	0	10	50	24.0
ラーメン	1	0	0	0	0	12	1	0	0	0	0	0	0	0	0	2	0	**32**	2	0	50	64.0
結婚式	0	0	0	1	0	9	5	0	0	0	0	0	0	0	2	7	0	0	**26**	0	50	52.0
槍ヶ岳	0	0	0	0	0	4	0	0	0	0	0	0	0	1	1	5	0	0	0	**39**	50	78.0
分類枚数	20	6	3	55	16	169	78	2	0	8	36	5	15	17	59	216	27	71	71	126	1000	**38.5**
適合率	90	66	66	47	81	18	48	100	0	100	83	100	86	94	49	18	44	45	36	30	**60.4**	**47.0**

表 9.15 実験 8：方法 3 による実験結果の confusion matrix

	りんご	クマ	マウン	琵琶湖	乗用車	ネコ	入学式	住宅	イチロ	観覧車	金閣寺	ライオ	モアイ	ノート	新幹線	公園	ペンギ	ラーメ	結婚式	槍ヶ岳	**枚数**	**再現率**
りんご	**21**	5	0	0	0	2	2	1	0	1	1	2	0	0	2	5	1	5	0	2	50	42.0
クマ	2	**13**	1	4	0	6	1	0	0	0	3	2	2	0	3	2	1	2	4	4	50	26.0
マウンテンバイク	2	3	**7**	3	1	4	2	0	0	3	4	0	0	1	4	8	4	0	3	1	50	14.0
琵琶湖	0	2	0	**28**	0	2	0	0	0	2	0	0	1	0	1	4	3	0	0	7	50	56.0
乗用車	0	3	0	2	**19**	4	1	2	1	0	3	0	0	2	6	2	2	1	2	0	50	38.0
ネコ	1	9	1	0	0	**20**	1	0	0	0	2	2	0	0	1	1	2	2	2	6	50	40.0
入学式	0	0	1	2	1	0	**39**	0	0	1	1	0	0	0	0	1	0	0	4	0	50	78.0
住宅	1	2	3	4	0	0	7	**11**	0	3	4	0	0	0	3	6	2	1	1	2	50	22.0
イチロー	0	3	3	2	1	1	6	0	**3**	2	5	2	1	1	4	5	6	3	2	0	50	6.0
観覧車	0	0	2	2	3	3	3	0	0	**20**	2	1	1	0	2	2	2	0	2	5	50	40.0
金閣寺	1	0	0	3	0	1	0	0	0	1	**37**	0	1	0	2	2	1	1	0	0	50	74.0
ライオン	1	3	1	2	0	20	1	0	1	1	3	**5**	0	0	1	3	0	4	2	2	50	10.0
モアイ	1	2	1	2	3	1	0	1	0	0	1	0	**25**	0	1	4	0	1	1	6	50	50.0
ノートパソコン	0	3	1	0	2	2	1	2	1	0	0	0	1	**23**	5	0	7	0	0	2	50	46.0
新幹線	0	2	0	2	0	1	3	0	0	2	1	0	0	0	**30**	1	2	3	3	0	50	60.0
公園	2	2	2	3	0	3	1	1	0	0	13	0	3	0	0	**14**	0	0	1	5	50	28.0
ペンギン	0	5	1	5	3	6	2	1	0	0	3	0	0	1	4	4	**9**	0	3	3	50	18.0
ラーメン	0	1	0	0	1	7	3	1	0	0	1	1	1	1	1	0	2	**28**	0	2	50	56.0
結婚式	1	4	0	2	0	1	8	0	0	0	2	0	0	0	5	2	2	5	**16**	2	50	32.0
槍ヶ岳	0	2	1	1	1	1	0	0	0	0	2	0	4	0	1	2	2	0	0	**33**	50	66.0
分類枚数	33	64	25	67	35	85	81	20	6	36	88	15	40	29	76	68	48	56	46	82	1000	**40.1**
適合率	63	20	28	41	54	23	48	55	50	55	42	33	62	79	39	20	18	50	34	40	**43.1**	**41.5**

表 9.16 実験 9：方法 3 による実験結果の confusion matrix

	りんご	クマ	マウン	琵琶湖	乗用車	ネコ	入学式	住宅	イチロ	観覧車	金閣寺	ライオ	モアイ	ノート	新幹線	公園	ペンギ	ラーメ	結婚式	槍ヶ岳	**枚数**	**再現率**
りんご	**18**	2	0	0	0	3	4	2	0	0	2	1	0	0	2	8	0	8	0	0	50	36.0
クマ	1	**9**	0	6	0	12	1	0	0	0	1	1	1	0	3	4	2	3	5	1	50	18.0
マウンテンバイク	1	0	**5**	2	3	5	5	0	0	2	0	0	0	0	2	15	3	2	5	0	50	10.0
琵琶湖	1	0	0	**33**	0	2	0	1	0	0	0	0	2	0	0	4	2	0	1	4	50	66.0
乗用車	0	1	4	0	**18**	1	3	2	0	0	2	0	1	1	9	2	1	2	3	0	50	36.0
ネコ	1	7	0	0	0	**25**	0	0	0	0	2	3	0	0	0	3	3	2	0	4	50	50.0
入学式	0	0	0	0	0	0	**44**	0	1	0	0	0	0	0	0	0	0	1	4	0	50	88.0
住宅	0	1	1	1	2	1	4	**15**	0	0	2	0	1	0	3	8	4	1	3	3	50	30.0
イチロー	0	0	1	1	1	1	9	0	**4**	0	1	2	0	0	6	11	1	4	7	1	50	8.0
観覧車	0	0	1	3	3	6	4	1	0	**13**	0	1	1	0	1	7	1	0	3	5	50	26.0
金閣寺	0	0	1	1	0	1	0	0	0	0	**32**	0	1	0	2	7	0	2	0	3	50	64.0
ライオン	1	3	0	1	0	25	1	0	1	0	0	**5**	0	0	0	4	0	6	3	0	50	10.0
モアイ	1	0	1	4	2	2	0	0	0	0	2	0	**22**	0	0	8	1	1	1	5	50	44.0
ノートパソコン	0	1	2	3	3	0	2	0	0	0	0	0	1	**21**	2	3	9	0	0	3	50	42.0
新幹線	0	4	0	0	1	3	2	0	0	0	1	0	0	0	**27**	2	2	2	6	0	50	54.0
公園	5	0	0	0	0	4	2	1	0	0	6	0	1	0	0	**29**	0	0	1	1	50	58.0
ペンギン	0	3	1	6	4	9	3	0	0	0	1	0	0	0	4	4	**9**	0	1	5	50	18.0
ラーメン	0	1	0	0	2	5	1	1	1	0	0	0	1	1	1	2	1	**30**	1	2	50	60.0
結婚式	2	1	0	1	0	2	7	0	0	0	0	0	0	0	5	4	1	6	**20**	1	50	40.0
槍ヶ岳	0	1	0	1	0	1	0	0	0	0	1	0	1	1	0	1	0	1	0	**42**	50	84.0
分類枚数	31	34	17	63	39	108	92	23	7	15	53	13	33	24	67	126	40	71	64	80	1000	**42.1**
適合率	58	26	29	52	46	23	47	65	57	86	60	38	66	87	40	23	22	42	31	52	**47.9**	**44.8**

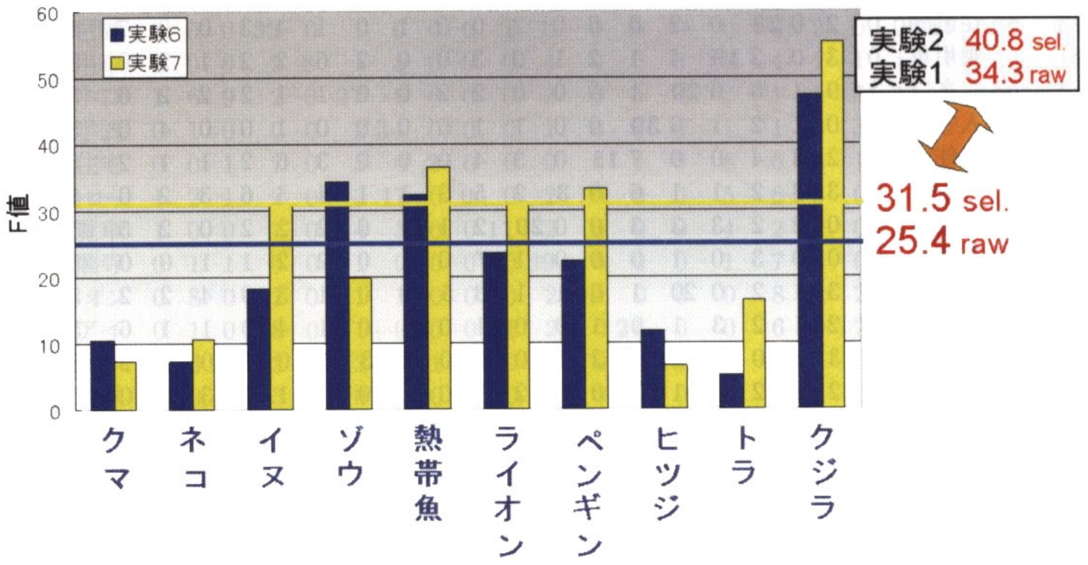

図 9.26 実験 6, 7 の結果 (方法 2).

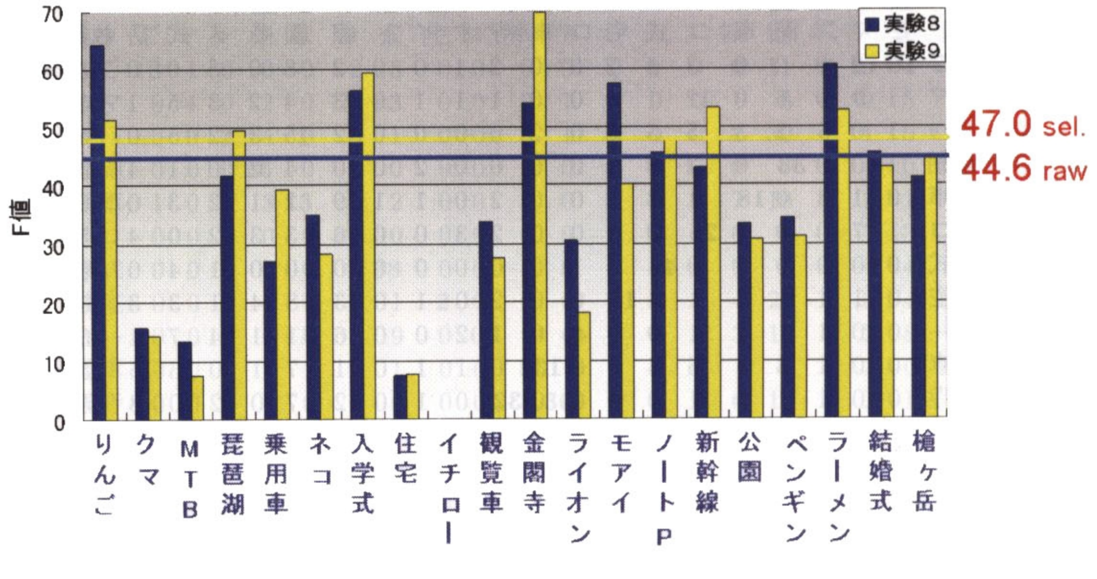

図 9.27 実験 8, 9 の結果 (方法 2).

9.4.3 再現率–適合率グラフ

本節では，実験 8 と実験 9 において分類時に距離に関する閾値を設けて，入力画像と距離最小のクラスとの距離が閾値以上の場合は「分類なし」，つまりリジェクトとすることとした．その上で，閾値を変化させて，横軸を再現率，縦軸を適合率として，「再現率–適合率グラフ」を作成した．なお，本章の分類実験においては，1 枚の入力画像を複数のクラスに同時に分類することは行わないので，再現率の上限は閾値を用いなかった場合の結果と等しくなる．再現率は閾値を 0 にすれば 0 になるので下限は 0 であるが，0.1 未満の場合は各クラスの正しく分類された画像が平均 5 枚未満となり，サンプル数が少な過ぎて，グラフの変化が非常に不規則であったため，グラフでの再現率の下限値は 0.1 とした．

図 9.28，図 9.29は，それぞれ，実験 8, 実験 9 において，距離の閾値を変化させた場合の方法 1 から方法 3 までの 3 種類の分類方法に関する「再現率–適合率グラフ」である．なお，方法 3 と方法 4 はほぼ同じ手法であるので，方法 4 については省略する．実験 8 では，方法 1 と方法 3 の場合は再現率 0.1 近くで適合率が最大になり，再現率が上昇するにつれ適合率が徐々に減少している．方法 2 の場合は再現率 0.15 で最大となり，次第に減少するものの，再現率 0.3 付近で 2 つめのピークが存在している．実験 9 では，実験 8 の場合とは異なり，徐々に適合率が上昇して，上限に近付くと適合率が減少している傾向がうかがえる．

ここでは，「画像特徴空間における距離が近い画像ほど，互いに同じクラスの画像である可能性が高い」ことをグラフから読取れることを期待していたが，その傾向は図 9.28，図 9.29にはあまり顕著には現れていない．そこで，次に，分類しにくい画像クラスを除外するために，実験 8，実験 9 での 3 つの方法での F 値の平均値の上位 10 クラスのみ (「りんご」「琵琶湖」「入学式」「観覧車」「金閣寺」「ノートパソコン」「新幹線」「公園」「ラーメン」「槍ヶ岳」) についての「再現率–適合率グラフ」を作成した．実験 8 の場合，実験 9 の場合をそれぞれ図 9.30，図 9.31に示す．図 9.30，図 9.31ともに，実験 8 の方法 2 を除いて，再現率 0.3 付近で最大の適合率を示し，再現率の増加と共に適合率が減少する傾向を示している．実験 8 の方法 2 については，再現率 0.2 がピークとなっている．ある一定の距離までは適合率が増加し，それ以降は減少する傾向があることから，分類し易い画像クラスにおいては「画像特徴空間における距離が近い画像ほど，互いに同じクラスの画像である可能性が高い」ことが言える．

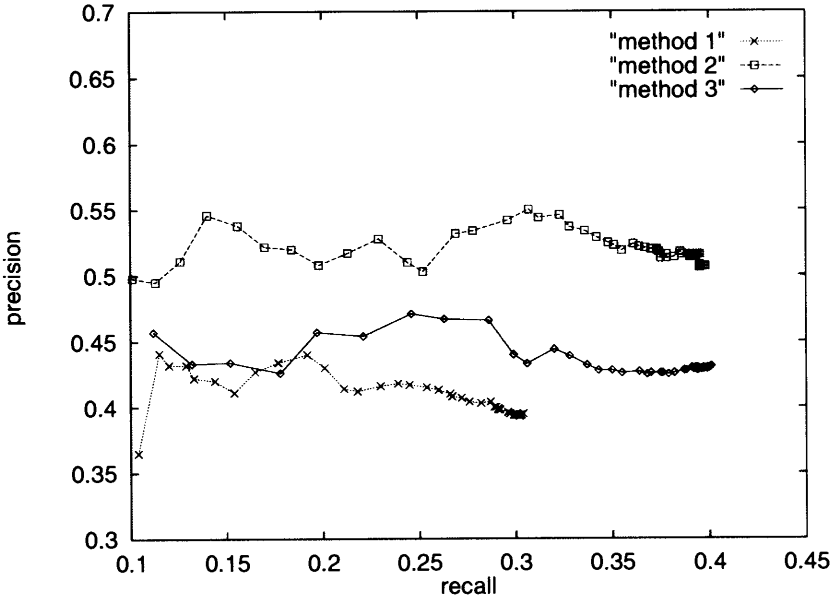

図 9.28 実験 8 ：再現率–適合率 (recall–precision) グラフ

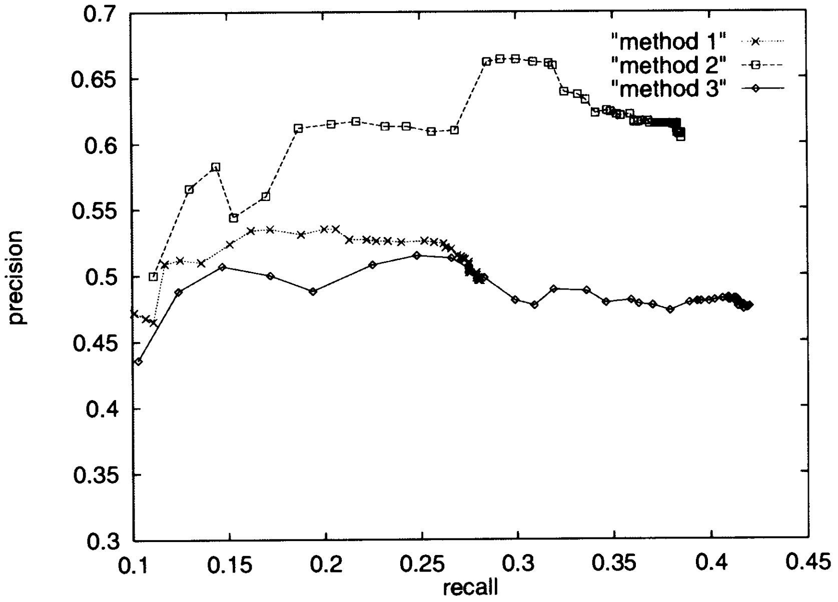

図 9.29 実験 9 ：再現率–適合率 (recall–precision) グラフ

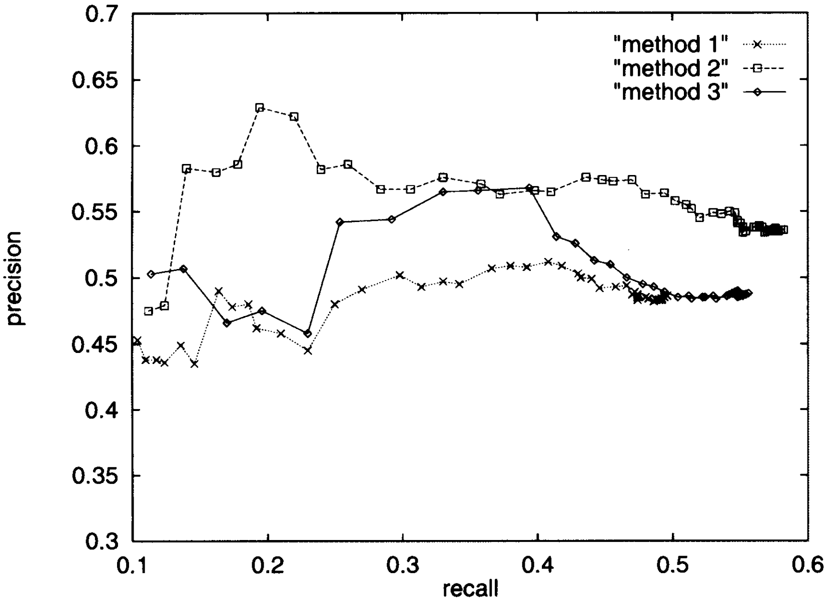

図 9.30 実験 8 ：F 値が平均以上のクラスのみの再現率–適合率 (recall–precision) グラフ

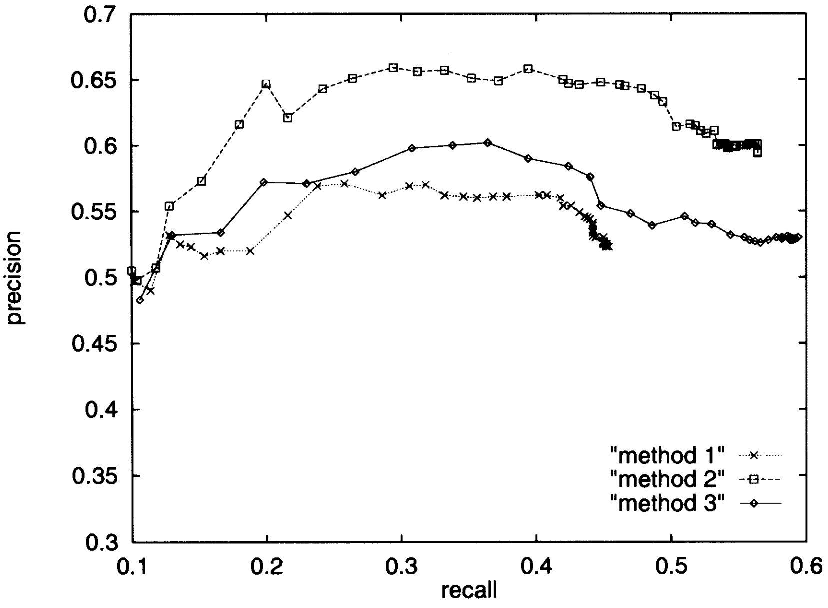

図 9.31 実験 9 ：F 値が平均以上のクラスのみの再現率–適合率 (recall–precision) グラフ

9.4.4 パラメータを変化させた場合

分類実験では 3 種類の分類方法を用いたが，それぞれに決めるべきパラメータが存在した．方法 1，方法 2 については画像をブロック分割する際のブロック数，方法 3 については k-NN 探索を行う場合の k の値をそれぞれ決める必要があった．本節では，分割ブロック数と k 値の変化による得られる F 値の値変化について行った実験について述べる．

表 9.32，図 9.17に方法 1 において分割ブロック数を 1, 5, 13, 25, 39 と変化させた時の実験 2 および実験 3 の結果を示す．なお，39 ブロックの時のみ，画像を 3 段階に縮小して，複数の解像度からブロックの切出しを行った．具体的には，画像を 240×180 の大きさに正規化し，0.7 倍の倍率で 3 段階に縮小することによって，3 つのレベルの多重解像度画像を用意し，それぞれのレベルのから半ブロックずらしたブロック画像も含めてそれぞれ 16+9, 9+4, 1 個 (合計 39 個) の 60×45 ブロック領域を抽出した．実験 2 では 39 ブロックの時が最も良くなり，実験 3 では 13 ブロックの時が最も良くなっている．このことから，分類実験では中間値の 25 ブロックを採用した．

また，方法 2 において，同様に分割ブロック数を 1, 5, 13, 25, 39 と変化させた時の実験 2，実験 3 の結果を表 9.33，図 9.18に示す．どちらの場合に置いても，39 ブロックの場合は最も良い F 値になっている．しかし，39 ブロックの場合と 25 ブロックの場合の差は僅かしかなく，処理時間がブロックに比例した時間掛ることを考え，分類実験では 25 ブロックを採用した．

表 9.17 方法 1 において，分割ブロック数を変化させた場合の分類結果 (F 値)．

分割ブロック数	1	5	13	25	39
実験 2 (10 クラス)	24.3	28.7	31.0	33.6	**33.8**
実験 3 (Corel 画像)	32.6	45.9	**49.1**	47.2	44.8

表 9.18 方法 2 において，分割ブロック数を変化させた場合の分類結果 (F 値)．

分割ブロック数	1	5	13	25	39
実験 2 (10 クラス)	25.0	32.2	35.6	38.1	**39.7**
実験 3 (Corel 画像)	38.7	49.5	57.0	59.0	**60.3**

次に，方法 3 において，k-NN 探索の k の値を 1, 3, 5, 10, 15, 20 と変化させた場合の実験 1，実験 2，実験 3 の結果を表 9.34，図 9.19に示す．実験 1 では k の値が 3 および 5 の場合，実験 2 では 5 の場合，実験 3 においては 3 の場合にそれぞれ最も F 値が良くなっている．そのため，分類実験に置いては，$k = 5$ を用いることとした．

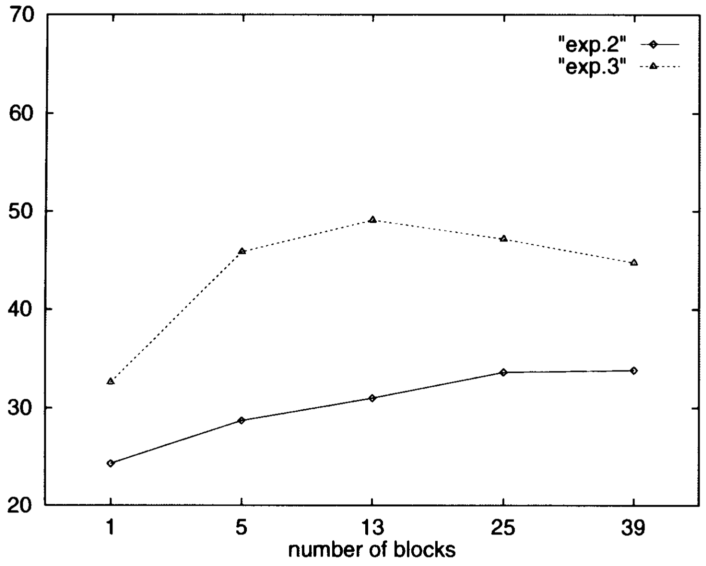

図 9.32 方法 1 において，分割ブロック数を変化させた場合の分類結果．分割ブロック数と F 値の関係を示す．

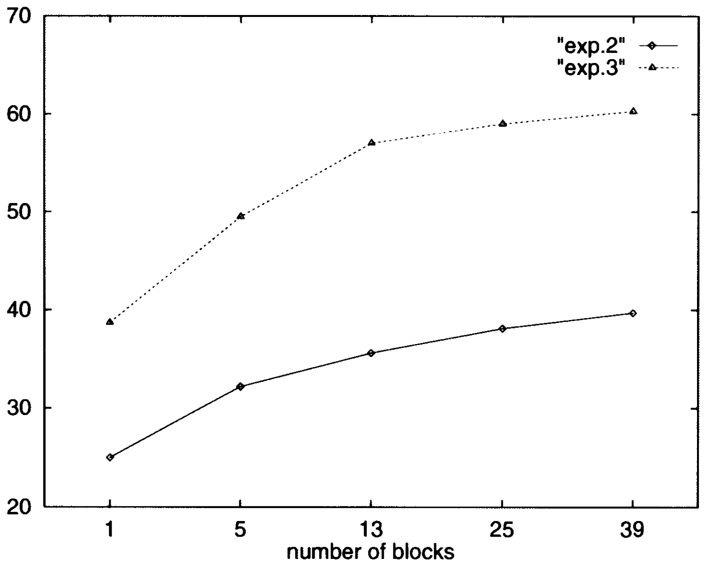

図 9.33 方法 2 において，分割ブロック数を変化させた場合の分類結果．分割ブロック数と F 値の関係を示す．

表 9.19 方法 3 の k-NN 法における k の値を変化させた場合の分類結果.

k の値	1	3	5	10	15	20
実験 1 (10 クラス)	30.3	30.8	**30.8**	28.3	26.8	25.9
実験 2 (10 クラス)	34.2	35.2	**36.0**	34.2	33.9	32.7
実験 3 (Corel 画像)	66.4	**68.9**	68.1	65.0	64.0	59.5

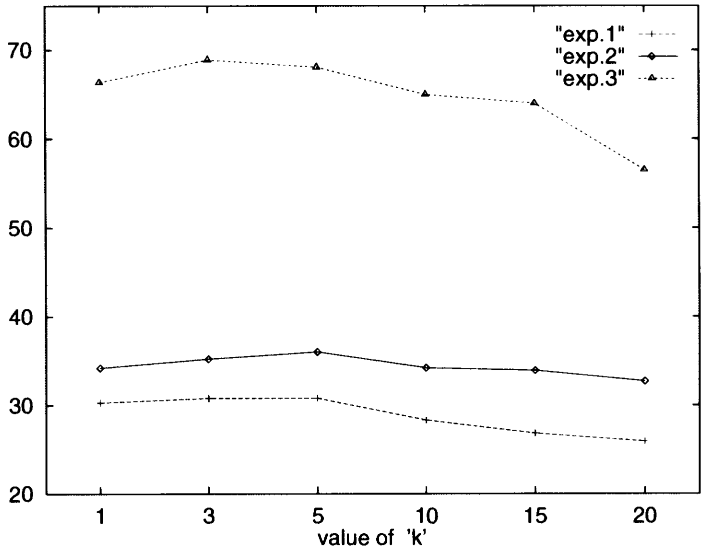

図 9.34 方法 3 の k-NN 法における k の値を変化させた場合の分類結果.

9.4.5 SR-tree による高速化

本章における画像分類では，基本的に画像特徴の類似検索に基づいて分類クラスを決定するので，学習画像が増えれば増えるほど分類に要する時間は増大する．特に工夫をしない場合は，通常は全探索を行うことになるので，学習画像の枚数に比例しただけ分類に時間が掛ってしまい，分類実験の規模を大きくする際の障害となってしまう．そこで，本研究においては，高速化のために高速インデクシング手法である SR-tree を導入することを試みた．そこで，本節では，高速インデクシング手法である SR-tree を利用した場合と通常の全探索の場合の実行時間について比較を示す．

ここでは分類実験の結果があまりよくなかった方法 1 を除いた方法 2，方法 3 の 2 つの分類方法について，多次元ベクトルの高速インデクシング法である SR-tree を用いて高速化を行った結果について述べる．

まず，方法 2 の場合に実験 2，実験 4，実験 5 の 3 つの分類実験を全探索で行った場合と SR-tree で行った場合の分類に要した時間の比較を表 9.20に示す．この分類時間は，交差法における 1 回の試行中に分類処理に要する時間であり，実際の分類実験ではこれを複数回行ってその合計を分類結

表 9.20 方法 2 の場合の全探索時と SR-tree 使用時の実行時間の比較．ブロック分割数は 25.

	全探索 (秒)	SR-tree (秒)	短縮倍率
実験 2 (10 クラス, 3052 枚, テスト画像全 200 枚)	1322.23	714.90	0.54
実験 3 (Corel 画像, 500 枚, テスト画像全 100 枚)	132.76	117.94	0.88
実験 5 (20 クラス, 3485 枚, テスト画像全 400 枚)	4812.50	3198.10	0.66

表 9.21 方法 2 の場合の全探索時と SR-tree 使用時の分類結果 (F 値) の比較.

	全探索	SR-tree	低下 F 値
実験 2	38.1	39.1	-1.0
実験 3	59.0	57.5	1.5
実験 5	43.3	39.9	3.4

果とした．また，分類に要する時間以外にも，学習画像とテスト画像の読み込みおよび特徴量抽出に要する時間があるが，これらの処理は予め 1 回行っておいてその結果を保存しておけば再利用できるので，ここでは考えないこととする.

方法 2 では 25 ものブロックを用いて 1 枚の画像を分類するので，例えば，実験 2 で 1 回の試行でテスト画像を 200 枚分類するには，実際には 5000 枚のブロック画像を分類していることとなる．そのため，実験 2 の全探索時には約 22 分もの実行時間が掛っている．SR-tree を用いると 6 割程度に減少するものの，まだ 12 分程度の時間が掛っている．実験 4 の 20 クラスの分類実験の場合は 80 分が 53 分に減少している.

次に表 9.21に全探索の場合と SR-tree の場合の得られる F 値の比較について示す．SR-tree は多次元ベクトル間のユークリッド距離に基づいて効率的に k-NN 探索を行うことが可能であるが，方法 2 の場合はベクトルと重み値の集合によって表現される signature として画像特徴を表現し，Earth Mover's Distance (EMD) によってその signature 間の距離を求めるので，そのままでは適用出来ない．そこで，signature 間の距離の下限値が signature の重心ベクトル間のユークリッド距離になることを利用して，下限値を SR-tree でインデクシングし，分類時にはテスト画像との距離の下限値の小さい順に 100 個程度の学習画像を取出して，それらとの間で実際の EMD を計算して，クラス分類を実現する．この方法では，実際には最近距離に存在する可能性のあるすべての候補との間で EMD を計算しているわけでないので，全探索と異なる結果になる可能性がある．そこで，ここでは全探索との結果の比較を行う．表 9.21によると，実験 2 を除いては，SR-tree 使用時の方が若干 F 値が減少していることが分かる．この低下は実際に EMD を求める個数を上位 100 個から増やすことによって防ぐことが出来るが，処理時間はその分増大する．つまり，精度と時間のトレードオフになる.

表 9.22は方法 3 における全探索時と SR-tree 使用時の実行時間の比較を示している．こちらでは,

表 9.22 方法 3 の場合の全探索時と SR-tree 使用時の実行時間の比較. $k = 5$.

	全探索 (秒)	SR-tree (秒)	短縮倍率
実験 2 (10 クラス, 3052 枚, テスト画像全 200 枚)	5.47	1.50	0.27
実験 3 (Corel 画像, 500 枚, テスト画像全 100 枚)	0.80	0.31	0.38
実験 5 (20 クラス, 3485 枚, テスト画像全 400 枚)	32.53	4.56	0.14

表 9.23 方法 3 の場合の全探索時と SR-tree 使用時の分類結果 (F 値) の比較. $k = 5$.

	全探索	SR-tree	低下 F 値
実験 2	36.0	35.7	0.3
実験 3	68.1	64.0	4.1
実験 5	38.9	34.7	4.2

最大 6 倍以上の高速化が達成できており，方法 2 の時と比べると大幅に時間短縮になっていることが分かる．表 9.23の F 値の変化であるが，SR-tree を用いることによって最大 1 割程度の低下が見られる．しかし，方法 2 の場合と同様に実際に EMD を計算する個数を増やすことで F 値を全探索時の値に近付けることが可能である．

9.5 まとめ

本研究では，WWW から収集した画像を学習画像として，画像認識 (画像分類) を行うことを試みた．従来，WWW からの知識獲得は，数値データやテキスト情報が中心であったが，本研究では画像情報についての WWW からの知識獲得を行ったことになる．従来の画像認識システムで行われていた，画像知識ベースを構築するための学習画像を人手で与えるのとは異なり，キーワードを入力するのみで，画像に関する知識やさらには学習画像自体もユーザが与えることなく，WWW から学習画像を自動的に収集して，実世界画像の自動分類を行うシステムを実現した．実験では，20 クラスの画像分類において，WWW から収集した画像をそのまま学習画像として用いた場合，最大 44.6%の F 値を得ることが出来た．

今後，分類精度をさらに向上させるための課題として，以下のようなことが挙げられる．

- より精度の高い分類手法の開発．
  - 今回は既に提案されている手法を用いて分類実験を行ったが，すべてが適切な画像あるとは限らない WWW 収集画像の特性を補うような画像分類方法を考案する必要がある．
  - 現在の k–NN 法では，これ以上画像枚数が増えた場合に対応することが難しい．予め学習画像をクラスタリングしておいて，クラスタの代表値を用いて分類することが考えられる．
  - Neural network(NN), support vector machine(SVM) などの他の分類器の利用．Support vector machine を用いた一般画像の分類 [203, 204] では，Corel 画像を用いた 14 クラスの分類において，85%程度の分類精度が報告されている．他の研究 [205, 206] でも SVM によって従来よりよい分類結果が報告されている．
- 学習画像となる WWW からの収集画像の適合率と収集枚数の向上．
  - 現在，適合率は 6〜7 割程度であるが，ノイズ画像がより少ない方が分類精度が向上することは実験より明らかである．そのため，収集画像の適合率向上が必要である．
  - WWW 画像収集システムにおける B 群画像選択の閾値を厳しくし，収集枚数を減らす代りに適合率を向上させる，WWW 画像収集システムでのクラスタリング処理に color signature と EMD を導入する，などが考えられる．
  - 収集枚数を各クラス 1000 枚以上にする．収集枚数が多い方がより多様な画像が学習画像に含まれ，多様な画像に対応した分類が可能になることが期待される．
- 画像特徴量の検討．
  - 画像特徴空間において，抽出した特徴量がクラス間で分離されていないと，そもそも分類は不可能である．分類精度を上げるには，画像特徴の検討も必要である．

- 対象の切出しを行う.
  - 本章の分類実験では，対象の切出しを行っていない．そのため，キーワードに関して収集した画像と分類クラスを一対一で直接対応づけて良いのかという問題がある.
  - 画像には複数の物体が含まれているのが普通である.「犬」と「猫」が同時に写っている画像も存在するので，領域分割が必要である．けれども，あらゆる画像を適切に処理する領域分割アルゴリズムは存在しないという問題がある.
  - 領域分割ではなく，画像同士の部分マッチングを行う．学習画像とテスト画像との類似度を画像全体で調べるのではなく，予め学習画像からその中で共通の要素を抽出してから，それとテスト画像の類似度を調べる方法も考えられる．部分マッチングの方法は従来より多くの研究がなされている [207].
- 「分類」ではなくて,「認識」するべきでないか？
  - 現在はすべての画像を予め定義されたクラスのどれか 1 つに必ず分類しているので,「認識」ではなく「分類」を行っているに過ぎない．そのため，実世界にある画像を多くをカバーできるくらいの数のクラスを用意しない限り，実用的な「認識」を実現することはできない.
  - 「認識」を行うためには，どこのクラスにも属さないという「分類なし」クラスつまりリジェクトを導入する必要がある．これについては，第 9.4.3節で単純に閾値を導入した場合の実験を行ったが，十分には機能していなかった.「認識」を実現するためには，分類手法を含めて再検討する必要がある.
- 一般的な画像特徴，分類方法のみでよいのか？
  - クラスによって分類に適した画像特徴，分類方法が異なることは明らかである．そこで，特定対象向きの画像特徴，分類方法も組合せて利用する方法が考えられる.
- 何を認識・分類すればよいのか？
  - 「動物園のライオン」の画像があった場合に,「動物園」と「ライオン」の両方のクラスに分類するべきか,「動物園のライオン」クラスを設けるべきか？ つまり，1 枚の画像は同時に様々なクラスに分類される可能性があるので，どのようなクラスを設けて，最大いくつのクラスに分類するべきかは難しい.
  - 「ライオン」画像は「動物」画像でもある．画像認識は，画像を言語と結びつける処理であると言えるので，言語の階層構造とも密接に関係している.「動物」画像と「人工物」画像を最初に分類して，次に「4 本足動物」と「鳥，魚」などを分類するというように，階層的に認識を行うことも考えられる．なお，本研究においては，言語階層における「基本認識レベル」において認識・分類を行うことを基本とした.

- 多くのクラス，多くの画像での実験．評価の方法．
  - クラス数はいくつにしたら実用化できるか？ 各クラスの画像枚数はどの程度必要？
  - 評価はどうするべきか？ 評価用データによって，結果が大きく変わってくる．しかし，実世界のあらゆる対象に関して，評価データを用意することがほとんど不可能である．当面は，数百種類から数千種類にしぼって，評価用画像を用意することが妥当であろう．
  - 実世界画像は，地域性とも関係してくる．例えば，日本の「入学式」とアメリカの「入学式」は違う．日本の「住宅」とアフリカの「住宅」も違う．認識には知識が必要なので，与える知識によって認識結果は異なってくる．どのような知識を与えるかは，どのような認識システムを実現するかというスタンスによって異なってくる．

以上，実用化するためには今後様々な問題を克服することが必要である．究極的には，辞書に出ている普通名詞の数だけのクラスをある程度の精度で分類できるようにすることが目標である．

さらに別の問題として，WWW 上の画像の著作権の問題が挙げられる．WWW 上に存在している画像は著作権が存在する画像が大部分である．そのため，WWW から自動収集した画像を研究目的で使用する場合に，どの程度までの利用なら認められるか，またシステムを実用化する場合にはどうなのか，などについては，検討する必要がある．本研究のように大量に画像を収集する場合は，収集したすべての画像について著作者に承諾を得ることは事実上不可能である．そもそも Web ページに著作者の連絡先が記述されていない場合も多い．

研究目的に限っての論文への掲載に関しては，著作権法第三十二条の規定により，研究目的の場合は公表された著作物の引用が著作者の承諾を得ることなしに行うことが認められているために，引用元つまり画像の URL を明記した場合には著作権法上は問題ないと考えられる．しかしながら，すべての画像の URL を示すことは大量の画像を扱う場合にはこれも不可能に近いので，引用元を示さない引用に関しては，法律上の議論の余地がある．

ただし，日本の著作権法が適用されるのは，国内にサーバーがある Web サイトに限定される．それ以外の国にある場合は，それぞれの国の法律が適用される．世界中から画像を収集する場合には，各国の法律に従う必要があるが，世界中で 200 以上の国がインターネットに接続している現在，そうすることは事実上不可能であり，現状ではグレーの状態のまま，WWW から収集した画像の利用を行うしかないと言える．