

クロマキーによる画像合成精度向上手法

背景

クロマキー技術とは、画像内の画素値を元に前景、背景の領域を切り分ける処理手法である。また、この技術を応用することで画像を合成させることが可能となる。例として天気図と解説者の姿を合成する場合が挙げられる。

図1にクロマキー技術による画像合成を行なった結果を示す。

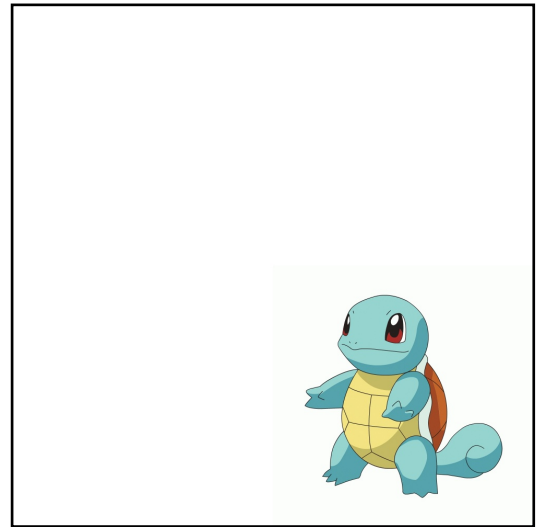


図1 背景のない画像におけるクロマキー合成結果

図1では、切り抜き対象の画像は背景が白いものである。そのため、物体との色の差を利用しやすく鮮明な合成画像を作成することができる。

図2に背景のある画像を対象にクロマキー合成を行なった結果を示す。



図2 背景のある画像におけるクロマキー合成結果

図2より、背景がある場合は前景と後景の差が小さくなり、物体の切り抜きがうまくできていないことがわかる。また、図3に切り分ける閾値を変えた出力結果を示す。

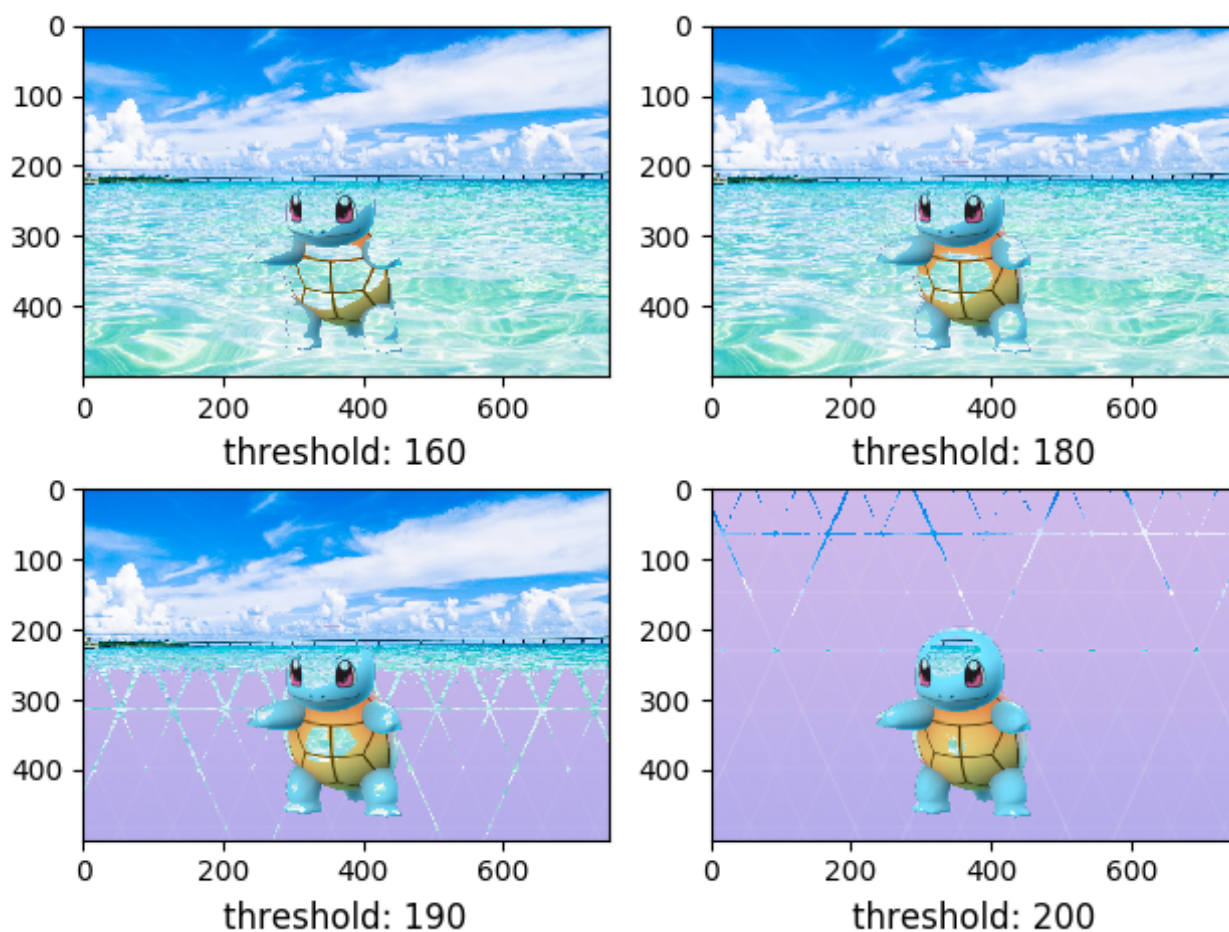


図3 前景、後景を切り分ける閾値と合成画像

図3より、今回使用する画像の場合は閾値180以下では対象物の切り抜きができなくなり、閾値180を超えた場合は背景除去ができなくなることがわかる。

よって今回は、クロマキー技術に関して背景のある画像への応用できるよう改良することを目的とする。

仮説

一般的なデジタル画像は、コンピュータの内部では輝度情報の集まりとして管理されている。また、カラー画像であれば赤、緑、青(RGB)の3つの輝度で表現されている。図4に画像と画素値の関係を示す。

104	133	155	157	145	128	103	73	57	52	52	48	28	21	98	185	143	64
104	133	155	157	145	124	86	49	26	24	24	22	12	17	98	185	144	100
104	133	155	157	144	125	94	62	42	39	39	35	22	18	98	185	158	112
52	67	77	78	73	61	54	54	53	52	50	50	36	22	77	169	160	81
52	67	77	78	73	59	41	28	25	24	23	24	14	13	77	169	169	106
52	67	77	78	73	60	48	41	40	39	38	37	26	16	79	169	170	119
31	33	32	33	35	31	37	50	53	51	51	51	40	21	61	181	172	101
31	33	32	33	35	29	26	24	25	23	24	24	18	15	61	181	178	120
31	33	32	33	35	30	32	37	40	38	39	39	30	19	63	151	178	124
27	27	27	27	27	27	32	46	50	52	52	52	46	24	48	134	178	123
27	27	27	27	27	27	27	24	23	24	24	24	20	13	46	133	182	131
27	27	27	27	27	27	31	36	38	39	39	39	33	19	47	133	183	140
27	27	27	27	27	27	26	43	49	52	51	50	49	25	35	118	181	144
27	27	27	27	27	27	27	27	26	25	24	23	23	12	35	114	181	154
27	27	27	27	27	27	29	36	39	39	38	38	36	19	35	115	183	156
27	27	26	27	28	27	28	41	48	51	50	50	49	31	27	99	178	161
27	27	26	27	28	27	26	24	24	23	23	23	22	11	22	99	178	169
27	27	26	27	28	27	32	37	38	38	38	37	22	26	97	180	171	
26	27	28	27	27	27	27	36	47	53	57	66	71	57	35	89	173	176
26	27	28	27	27	27	27	23	25	22	24	28	29	18	20	88	173	182
26	27	28	27	27	27	27	30	37	38	41	49	49	36	27	88	173	182
27	27	27	27	27	27	27	38	61	82	101	117	130	122	70	88	168	187
27	27	27	27	27	27	27	25	27	32	38	43	45	37	25	79	168	189
27	27	27	27	27	27	27	32	44	57	68	80	86	78	45	86	168	188
27	26	27	27	27	26	25	40	94	100	119	162	162	161	92	88	160	188
27	26	27	27	27	26	27	24	35	53	63	89	74	63	40	77	162	190
27	26	27	27	27	26	26	34	63	98	115	121	124	110	63	63	161	189
27	27	27	26	25	26	30	43	106	107	188	192	191	172	104	89	155	183
27	27	27	26	25	26	26	26	39	63	84	89	88	74	45	74	155	185
27	27	27	26	25	26	27	36	72	112	137	142	139	123	73	81	155	184
26	26	26	26	30	30	30	48	121	125	169	191	191	172	108	88	149	177
26	26	26	26	28	28	26	28	44	89	99	91	98	80	49	71	149	178
26	26	26	26	29	29	27	39	80	120	141	143	139	127	79	79	149	180
26	26	29	34	42	42	43	63	142	178	187	188	187	167	108	84	143	170
26	26	27	32	38	37	34	32	50	76	88	91	91	80	49	67	133	174
26	26	28	33	39	41	37	47	81	126	140	142	139	125	81	75	143	175
27	28	39	53	60	63	63	89	157	165	168	162	161	171	107	81	138	166
27	26	37	48	54	57	54	47	58	81	98	98	98	79	48	66	138	170
27	27	38	52	58	61	59	67	104	132	140	138	135	126	78	73	139	171
26	34	54	63	64	66	79	127	171	183	185	184	181	169	101	76	137	164
26	32	49	57	55	55	52	64	78	87	88	88	88	77	46	63	139	168
26	33	53	61	60	61	67	83	122	135	137	136	135	124	79	70	138	169
28	42	61	64	65	65	91	163	181	184	181	179	178	165	88	75	143	166
26	40	55	55	54	52	54	71	90	97	98	87	88	73	40	64	143	170
27	43	58	60	60	59	72	107	133	137	136	134	133	120	64	70	140	171

図4 画像と画素値の関係

通常のクロマキー合成では、3つの輝度をグレースケール変換によりまとめることで単一の輝度に変換しマスク画像の生成を行なってる。

しかし、各画素値で分けてマスク画像の生成を行うことで、より柔軟な合成画像の生成を行うことができるのではないかと考えた。

方法

1. 画素値ごとに画像を分ける
2. 1で作成した画像に対して二値化を行う。
3. RGBそれぞれの二値化画像を生成する。
4. 各二値化画像を合成しマスク画像を生成する。
5. 合成したマスク画像を使用し物体画像から対象物を切り抜く。
6. 合成したマスク画像を使用し背景画像から対象物のシルエットを切り抜く。
7. 方法5、方法6で作成した画像を合わせる。

上記の処理を順に行なっていくことで鮮明な合成画像を作成する。

使用する画像に関して、条件を同じにするため図2で示した物体画像と背景画像を使用する。図5、6に各画像を示す。



図5 背景画像



図6 物体画像

結果

図7に今回提案した方法により合成した画像を示す。



図7 今回提案した方法により合成した画像

図7より、輝度をまとめて合成を行う場合より綺麗な合成画像を作成することができる。

考察

今回、画素値について各要素ごとに分けて画像の前処理を行い、それぞれの画像を合わせてマスク画像を作成することにより鮮明な合成画像を作成できることがわかった。図8~10に各画素値における二値化画像を示す。

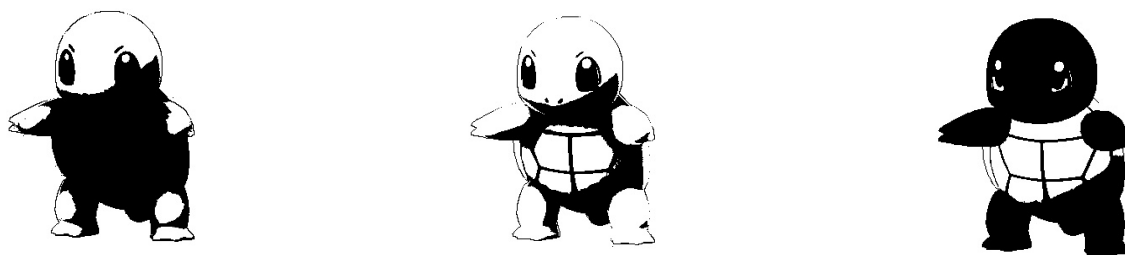


図8 各画素値における二値化画像(左: 赤色のみ, 中央: 緑色のみ, 右: 青色のみ)

図8より異なる範囲が黒色になっていることがわかる。各画素値で分けることで切り抜き対象物の異なった範囲をカバーすることができ、その結果鮮明なマスク画像の生成と合成画像の作成ができるようになったと考えられる。

展望

各画素値を二値化する際に用いる閾値について、手作業で探す手法をとっている。そのため、1つの合成画像を作成するだけで多くの時間を費やしてしまう。

上記課題について、要素ごとに分けた画像を学習データ、それに対する最適な閾値をラベルとしてそれぞれ機械学習を行い予測できるようにすることでより利用しやすい技術となると考えられる。

まとめ

今回、クロマキー技術を改良した手法を提案した。画像の要素ごとに分けて処理を行うことで精度が向上することがわかった。しかし、現状では手作業による閾値決定を行っているため、画像ごと3つの閾値の最適値を調べる必要がある。今後、これらの閾値の自動化を実現することでより多様な使い方ができるようになると考えられる。

～実務上の技術応用例～

画像から文字を認識するOCRでは、多くの場合背景に色が含まれている場合文字認識の精度が下がることがある。図9にOCR対象画像とtesseractにより文字認識をさせた結果を示す。



図9 OCR対象画像(左)とOCR結果(右)

図9より文字の認識ができていないことがわかる。図10にクロマキー技術により文字と白画像を合成させて背景除去を行なった場合の画像とtesseractによるOCR結果を示す。



図10 クロマキー技術による背景除去画像(右)とOCR結果(左)

図10より、背景除去をすることで文字認識の精度が向上したことがわかる。