

第 10 章 画像・データ・フォーマット

1. 画像フォーマット

画像を記録するためのフォーマット(書式)は数多く提案されてきた。それらのフォーマットの中で、特筆すべきフォーマットは GIF(ジフあるいはギフと発音)フォーマットである。当初、GIF フォーマットは多くの画像に利用されて来たが、あるとき、GIF フォーマットの利用費用徴収を GIF フォーマットの権利所有者が宣言した。実際に GIF フォーマットの利用費用の請求があったかどうか確かではないが、一斉に GIF フォーマットの利用が皆無となった。出力画像フォーマットの書式から、GIF フォーマットが次々と削除されて行った。この世界で、有料の画像フォーマットは流通しない。

(1) J P E G

JPEG は、JPEG 委員会が維持管理するフォーマットであり、www.jpeg.org においてその詳細が紹介される。これまでの JPEG は離散コサイン変換 (Discrete Cosine Transformation) により画像を圧縮してきたが、この欠点を補うように JPEG-2000 が発表された。JPEG-2000 は、圧縮のためにウェーブレット変換を採用し、これまでの JPEG が持っていた歪エラーを解消した。

J P E G 画像を作成する際に、画像の圧縮設定が可能である。画像の圧縮は、非可逆的圧縮であり、元に戻すことができない。また、画像の圧縮率を高くすると、画質が低下する。課題において、画質の低下を体験する。

(2) T I F F

T I F F (ティフ) は、タグ付き画像ファイルフォーマット (T a g g e d Image File Format) と呼ばれ、A d o b e 社が著作権を所有する画像フォーマットである。高次元のカラー画像を取り扱うことが可能である。特色としては、ファイルのヘッダー内に、画像サイズ、データ並び、画像圧縮などの基礎情報を含むタグ (Tags : 荷札、標識) を持つことである。TIFF は画像データを収納可能であるとともに、ベクトル情報による切り抜き領域情報も収納可能である。TIFF フォーマットは、データ欠損がないため、画像を提供するフォーマットとして採用されてきた。

TIFF フォーマットは、LZW (Lemple-Ziv-Welch) の可逆的圧縮のオプションを持つ。TIFF 画像表示の際にエラーを経験することがある。ほとんどの場合、LZW の圧縮があるため、読み込みエラーを発生する。TIFF フォーマットで保存する際に、LZW 圧縮を解除する。

(3) GeoTIFF

GeoTIFF (ジオ・ティフ) は、TIFFフォーマットに地理参照情報を組み込むパブリック・ドメインのメタデータ記述基準である。TIFF画像を地理情報システムの空間へ投影するための情報として、空間の座標系の定義 (距離あるいは緯度経度)、座標系の大きさ (投影するためのスケール)、座標系の種類、データの種類などの情報をTIFFフォーマットへ付加する。GeoTIFFを作成あるいは読み取るためのライブラリがGDALとして公表されている。

GeoTIFFは、TIFFフォーマットと干渉しないように設計されており、TIFFフォーマットのデータをハンドリング可能なソフトウェアであれば、問題なく表示可能である。

(4) PNG

ポータブル・ネットワーク・グラフィック (PNG) フォーマットは、著作権料を求めるGIFフォーマットにとって代わるため設計された。特に、WEBにおける画像取扱を支援することを目的として普及してきた。PNGフォーマットに関する情報は次のサイトにおいて閲覧可能である。

<http://www.libpng.org/pub/png/>

PNGフォーマットは、GIFと比較すると、表4-5のようになる。GIFの持つアニメーション機能をサポートしないが、その他の機能では、GIFより優れる。

表 4-5 PNGとGIFの比較

	PNG	GIF
アルファ・チャンネル (可変透過機能)	優れる	劣る
ガンマー補正 (明るさの補正)	優れる	劣る
2次元インターレース (間引き) プログレッシブ (進行型) 表示	優れる ・ 概念はGIF ・ 形式は2次元でJPEG	劣る
圧縮率	若干優れる (5 から 25%)	若干劣る 画質の低下なし

	画質の低下なし	
多重画像支援(アニメーション)	機能なし(単一画像のみサポート)	機能あり
支援形式	3色混合カラー画像 白黒画像 8ビットカラーテーブル画像	8ビットカラーテーブル画像のみ (カラーと白黒)

2. データフォーマット

(1) HDFフォーマット

階層的データフォーマット (Hierarchical Data Format : HDF)は、米国イリノイ州立大学の国立スーパーコンピュータ応用センター (National Center for Supercomputing Applications : NCSA、<http://hdf.ncsa.uiuc.edu/index.html>) が開発した科学データの運用のためのデータフォーマットである。HDF ライブラリーとし、データの入出力プログラム、解析ツール、可視化ツール、フォーマット変換ツールを含む。HDF 4. x (HDF 4 以前を含む) と HDF5 の二つのフォーマットが存在するが、フォーマットが異なり、互換性がない。HDF ライブラリーは、無料で利用可能であり、世界各国で環境科学、中性子研究、非破壊検査、航空宇宙研究などに利用されている。HDF は、NASA の HDF-EOS プロジェクト、DOE の改良型シミュレーションと計算プログラムに利用されている。

(2) HDF-EOS フォーマット

NASA は、1993 年に地球観測システム (Earth Observing System : EOS) のデータ記録方式として、NCSA の HDF フォーマットを、標準のフォーマットとして採用した。この地球観測システムは、衛星搭載のセンサーが収集するデータから構成され、地球変動研究プログラム (Global Change Research Program) を支援するものである。NASA の採用決定以来、NCSA は地球観測システムの運用のために膨大なプログラム開発を行った。地球観測システムのデータを取り扱う HDF-EOS を開発した。1999 年に打ち上げられたテラ (Terra) が地球観測システムの第 1 号衛星となり、HDF-EOS が採用され、2000 年 4 月に HDF-EOS に対応したデータの配布が始まった。引き続き、アクア (Aqua) が 2001 年に、オーラ (Aura) が 2002 年に打ち上げられた。

地球観測システムは、多くのセンサーから 1 日当たりテラバイトあるいはそれ以上のデータを配信する。HDF フォーマットとソフトウェアには、これらのデータを取り扱うための品質と自由度が必要とされる。約 1000 名のプロジェクトサイエンティストが第一線のユーザーとなり、さらに高次データを利用する GIS 応用から学生まで 30,000 名を越えるユーザが考えられる。

NASA が地球観測システムへ HDF を採用したことにより、他の観測計画あるいは外国宇宙機関による観測計画への HDF 採用が進んだ。1999 年春に打ち上げられたランドサット 7 号も HDF を採用した。

(3) その他のデータフォーマット

HDF フォーマットが提案される以前から、リモートセンシングデータの提供が行われており、提供者ごとに独自のフォーマットが提案され、利用されてきた。これらのフォーマットに共通するキーワードを説明する。HDF においては、以下の概念はフォーマットの中に記述されている。

(i) ヘッダー

データのライン数、各ラインごとのピクセル数、観測日などを記述した部分である。各データ提供者が独自にヘッダーの書式(フォーマット)を提案する。書式説明書に基づき、ヘッダーを読み取る。

(ii) ライン

ラインは、画像を構成する要素であり、走査鏡を持つセンサーでは、走査方向がラインとして扱われる。また、CCD などの固定センサーにおいては、飛行方向に直角の方向のデータ配列をラインとして扱う。

(iii) ピクセル、カラム

ラインを構成する一つ一つの画素に対応する要素である。センサーによりビット数が異なる。8 ビットの 1 バイトあるいは 16 ビットの 2 バイトが主流であり、容量を抑えるため整数で与えられる。一部には、実数を利用するデータ系もある。

(iv) ビット

ビット配列は、最上位と最下位のビットの順番の異なる 2 系列のデータが存在する。コンピュータのデザインにより、ビットの配列が異なるもので、多くの画像処理あるいは表示ソフトウェアは、ビット配列を反転する機能を持つ。

(v) バンド配列

多波長バンドのセンサーがほとんどであり、各バンドの配列が異なり、次の概念が存在する。

- ・ バンド・インターリーブ・バイ・ピクセル(BIP)

隣り合ったピクセルにバンドの順番ごとにデータを並べる方法である。図 4-44a に示すように、各ピクセルが一つ一つのマス目に対応し、右方向に順にカラム 1 とライン 1 のピクセルに対して、バンド 1 からバンド 5 まで順番に並び、続いてカラム 2 とライン 1 のピクセルの各バンドと並ぶ。

- ・ バンド・インターリーブ・バイ・ライン

一つのラインに一つのバンドのピクセルを配列し、続くラインに次のバンドのピクセルを並べる方法である。図 4-44b に示すように、右方向にバンド 1 のカラム 1 からカラム n まで並びラインを単位とし、バンド 2 のライン 1、バンド 3 のライン 1 と並ぶ。

- ・ バンド・シーケンシャル

一つのバンドのピクセルからなる。バンド1について、コラム1からn、ライン1からラインnまで並び一つのバンドとしてグループ化される。さらに、バンド2からバンドnまでバンド単位の配列である。

このように、ピクセル、ライン、バンド単位のデータ配列が提案され、利用されてきた。エンドユーザーは、データ提供者の記述するデータ書式記述書に基づき、提供されるデータを読み取る。このようなデータをプログラムを作成し、直接読み取る機会は少なくなってきた。

例えば、マルチスペックでは、エンドユーザが指定するデータの配列定義に従って、データを読み取る機能を提供する。バンドの並び配列とともに、データの前に配置されるヘッダー（先頭レコード）や、各ライン単位に配置されるライン番号などのバイト数を指定すると、その約束に従って、画像データを読み取る機能である。

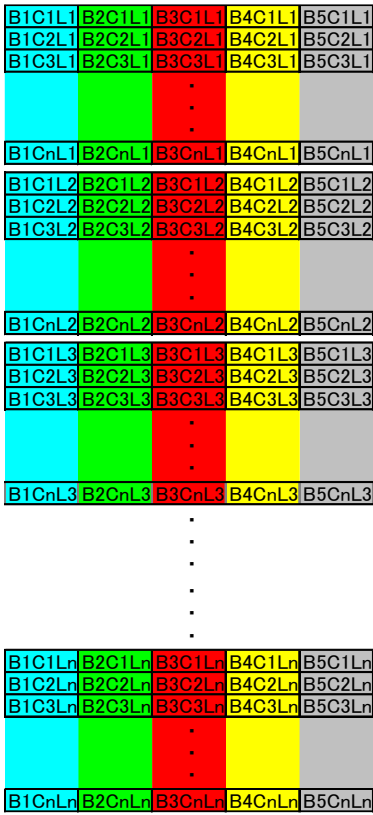


図 4-44a バンド・インターリーブド・バイ・ピクセル (BIP)
ピクセル (画素) 単位

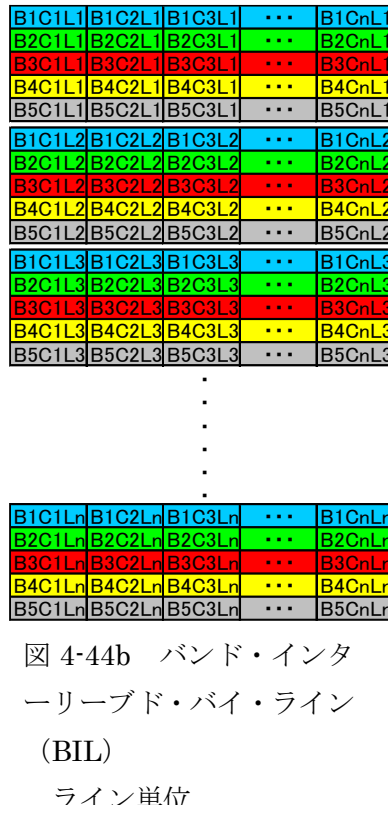


図 4-44b バンド・インターリーブド・バイ・ライン (BIL)
ライン単位

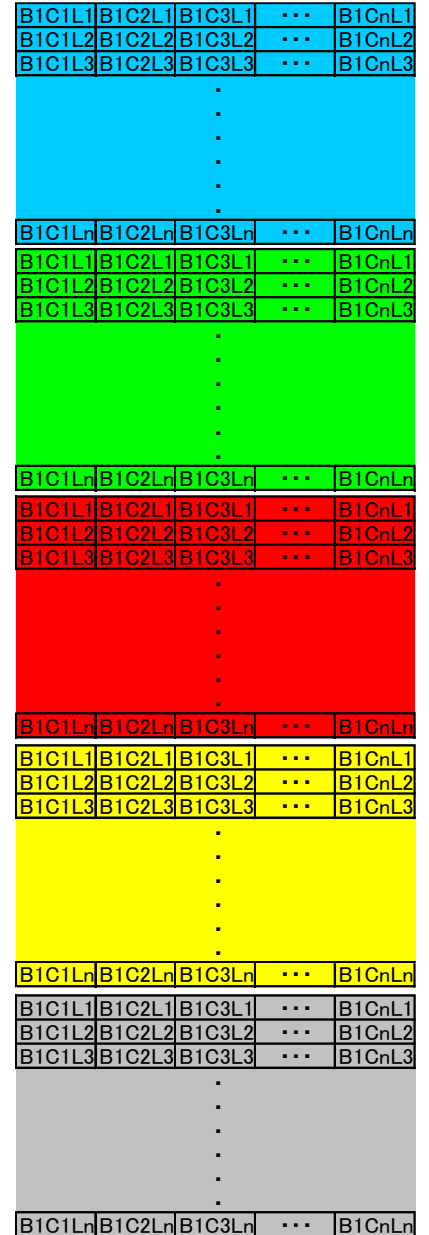


図 4-44c バンド・シーケンシャル (BSQ)
バンド(波長)単位

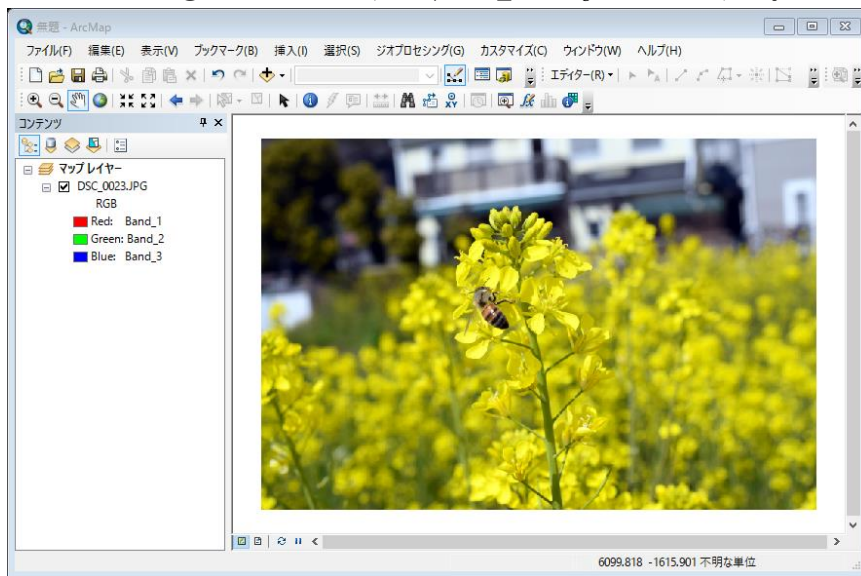
課題 10

JPEGファイルへ出力する際に、圧縮率を変化し、保存し、次の設問に答えよ。

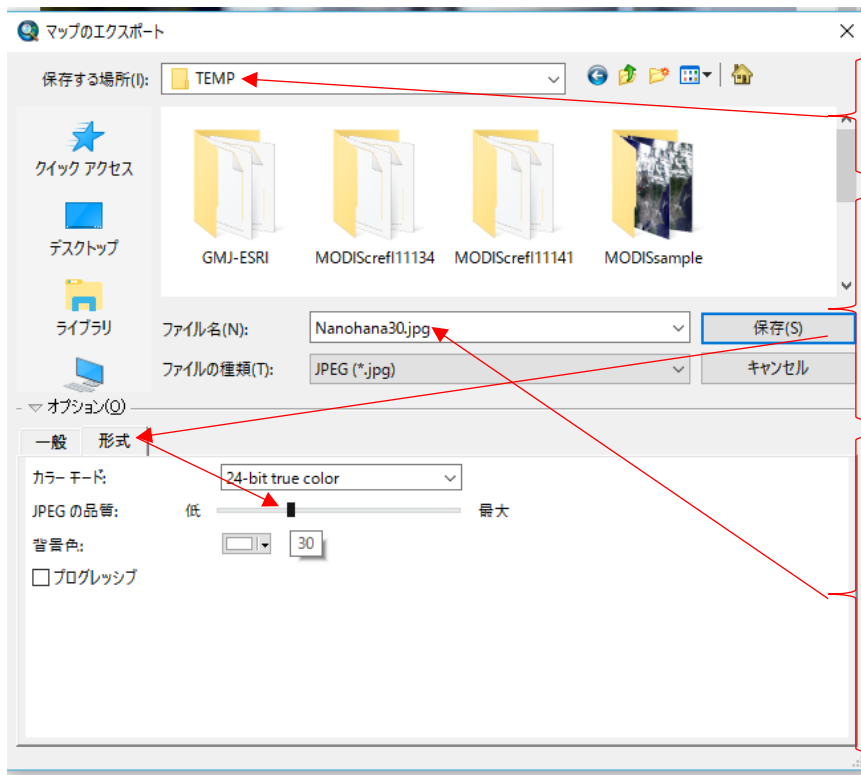
圧縮比較データ JPEGsample.zip を D:ドライブの¥TEMP フォルダへコピーし、解凍する。

- a. 圧縮率を100%、50%、20%、10%、5%、1%と変化させたときの保存ファイルのサイズを比較せよ。
- b. 同じく画質を主観的に比較せよ。

① ArcMAPを開き、DSC_0021.JPGを追加する。

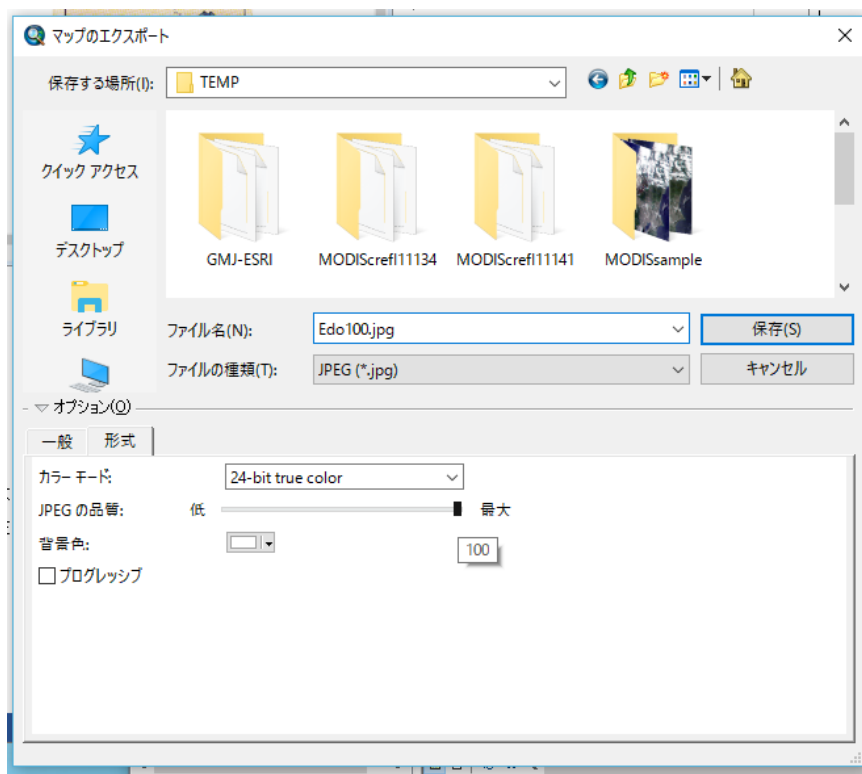
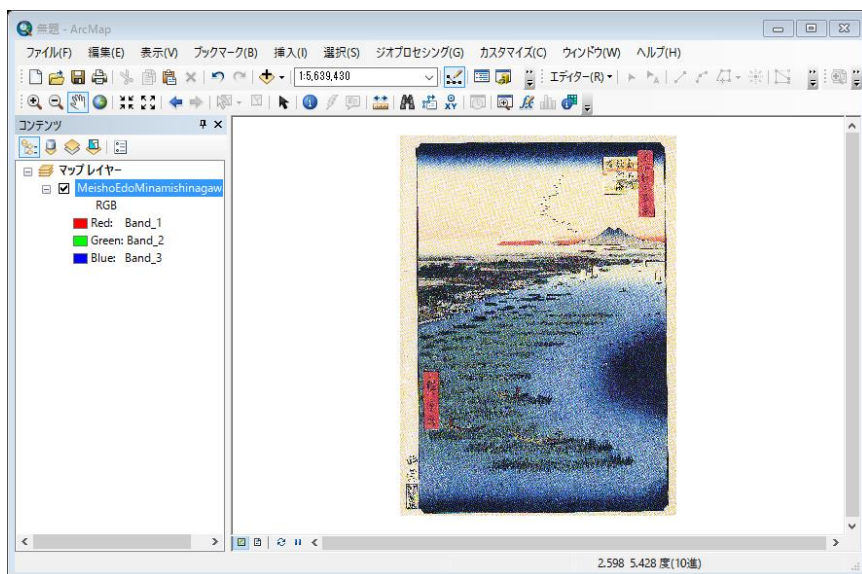


「ファイル」から「マップのエクスポート」を選択する。



出力先は、D:ドライブの¥TEMP フォルダとし、JPEGの品質を、形式タブのスライダーを操作し、100%、50%、20%、10%、5%、1%と設定の上、Nanohana100.jpg
Nanohana50.jpg
Nanohana20.jpg
Nanohana10.jpg
Nanohana5.jpg
Nanohana1.jpg
のように保存する。

- ② MeishoEdoMianamishinagawa.tif を追加する。広重の名所江戸百景南品川の浮世絵である。



前段と同様に、
JPEG の品質を設定し、
Edo100. jpg
Edo50. jpg
Edo20. jpg
Edo10. jpg
Edo5. jpg
Edo1. jpg
と保存する。

課題 10 JPEG の圧縮による画質の低下

学籍番号

氏名



品質 100% サイズ 2600KB



10% ○○○KB



50% ○○○KB



5% ○○○KB

みほん

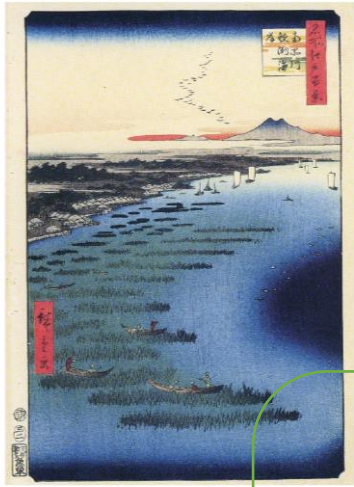


20% ○○○KB

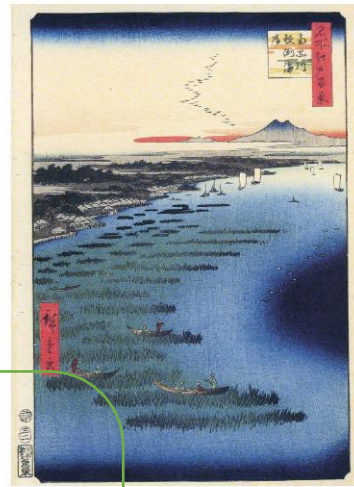


1% ○○○KB

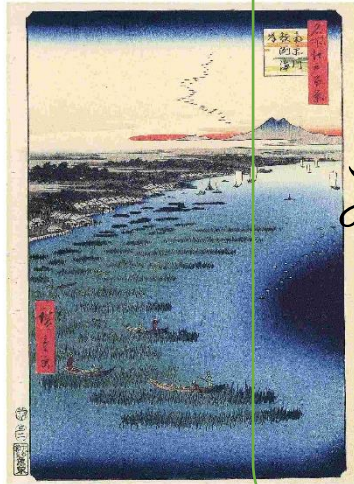
圧縮に従い 100%の 2600KB から 1%の○○○KB まで明らかにサイズが小さくなる。菜の花の画像の場合、○○%から被写界深度以遠の背景部の色変化が階段状となる。



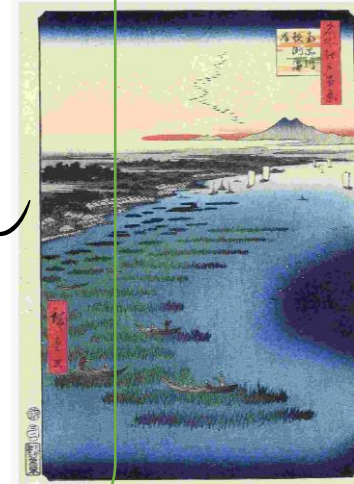
品質
100%
サイズ
3,700KB



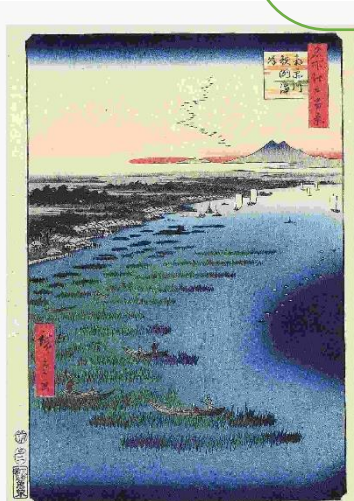
10%
〇〇〇KB



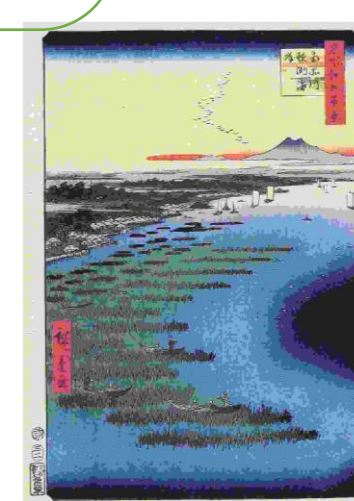
50%
〇〇〇KB



5%
〇〇〇KB



20%
〇〇〇KB



1%
〇〇〇KB

みほん

圧縮に従い 100%の 3700KB から 1%の〇〇〇KB まで明らかにサイズが小さくなる。浮世絵の場合、〇〇%から連続階調部の海の色変化が階段状となる。〇〇%ではノリ養殖の様子も画質の低下がみられる。