

Marcin KOCOT, Jacek GRUBER, Ireneusz J. JÓŹWIAK
Wydział Informatyki i Zarządzania
Politechnika Wroclawska

ZASTOSOWANIE MODELU ISINGA DO ODSZUMIANIA ZNISZCZONYCH DOKUMENTÓW

Streszczenie. W artykule przedstawiono sposób wykorzystania modelu Isinga w kontekście zastosowania go do usuwania szumów z dokumentów mających ubytki spowodowane uszkodzeniami fizycznymi. Proponowana metoda opiera się na maksymalizowaniu prawdopodobieństwa rozkładu łącznego za pomocą algorytmu iteracyjnej zmiany mody warunkowej. Zastosowane rozszerzenie algorytmu ICM pozwala na zwiększenie efektywności przetwarzania dokumentów. Wyniki badań empirycznych potwierdzają przydatność algorytmu w zastosowaniach.

Słowa kluczowe: odszumianie dokumentu, model Isinga, efektywność przetwarzania dokumentu.

APPLICATION OF ISING MODEL TO DAMAGED DOCUMENT DENOISING

Summary. This article discusses how to use the Ising model in the context of noise reduction being applied to the documents of losses caused physical harm. The proposed method is based on energy function minimizing using optimized Iterated Conditional Modes. Applied extension ICM algorithm allows to increase the efficiency of document processing. Empirical studies confirm the usefulness of the algorithm in practice.

Keywords: document denoising, Ising model, efficiency of document processing.

1. Wprowadzenie

Dwie główne kategorie zakłóceń w obrazie to rozmycie i zaszumienie. Rozmycie jest powiązane głównie z przetwarzaniem obrazu analogowego na postać cyfrową. Wynika to z faktu, że obraz cyfrowy może zawierać jedynie skończoną liczbę elementów [13]. Na

potrzeby tego artykułu przyjęto, że szum w obrazie to miejscowe zniekształcenia powodujące niechciane zmiany. *Odszumianie obrazów* (ang. *image denoising*) i ich restauracja jest jednym z podstawowych problemów przetwarzania obrazów (ang. *image analysis*). Głównym celem eliminacji szumów jest usunięcie zakłóceń i zniekształceń, które powstają na obrazie, przy jednoczesnym zachowaniu wszelkich jego istotnych cech, takich jak drobne detale, krawędzie czy też wzorce.

Szumy można podzielić na dwie kategorie. Pierwszą z nich jest fizyczne uszkodzenie, np. naderwanie lub podobny sposób trwałego naruszenia oryginału, blaknięcie atramentu bądź innego środka, którym utrwalono treść, przypadkowe lub celowe zabrudzenie obrazu za pomocą farby lub innej substancji. Drugą kategorią szumów są cyfrowe zniekształcenia wprowadzane wskutek transformacji analogowej postaci obrazu do postaci cyfrowej. Praca koncentruje się na obrazach dokumentów. Obrazy dokumentów charakteryzują się zawartością z dużą ilością tekstu i zawierają często rozmaite tabele oraz zwykle są czarno-białe. Proces odszumiania tego typu danych jest istotny, gdyż pozwala na odtworzenie zniszczonych fragmentów dokumentów istniejących jedynie w postaci drukowanej. Innym zastosowaniem usuwania szumów jest proces cyfryzacji dokumentów, szczególnie istotny w archiwizacji.

2. Ważniejsze metody odszumiania obrazów

Wyniki publikowane w literaturze przedmiotu koncentrują się na problemie eliminacji szumów z obrazów mających szum sklasyfikowany jako *sól i pieprz* (ang. *salt-and-pepper*) lub na innych, sztucznie dodawanych do obrazów zakłóceń. Mało jest opracowań i publikacji odnoszących się do odszumiania obrazów z zaszumieniami fizycznymi i zniszczeniami jego fragmentów lub do obrazów znacznie zniekształconych. Najbardziej znaczące w tej dziedzinie są metody klasyczne, np. filtry medianowe i uśredniające oraz metody oparte na transformacji falkowej.

Z puli algorytmów możliwych do zastosowania wybrano konkretne ich implementacje. Nacisk położono na dobór takich algorytmów referencyjnych, które pokrywały możliwie różnorodne podejścia do eliminacji szumów obrazów. Wśród rozpatrywanych algorytmów znalazły się algorytmy oparte na podejściu klasycznym, ale też nowe metody, hybrydowego odszumiania za pomocą filtrów Wienera i transformacji falkowej czy metody oparte na cząstkowych równaniach różniczkowych.

Dokonano obszernej analizy i wzięto do badań porównawczych następujące algorytmy, zacytowane w oryginalnym brzmieniu: algorytm De-Noise Color or Gray level images by using Hybrid DWT with Wiener filter (DWT) [11], Image Denoising using Fourth Order PDE (PDE) [15], algorytm Image and video denoising by sparse 3D transform-domain

collaborative filtering (BM3D) [7], Non-Local Mean Image Denoising (NLTV) [8], algorytm Nonlinear total variation based noise removal algorithms (ROF) [12], Image denoising with an optimal threshold and neighbouring window (SURE) [9], algorytm deconvtv fast algorithm for total variation deconvolution (TV) [5], filtr medianowy (MED), filtry wiener (WIE), filtr uśredniający (AVG). Jeśli algorytm wymagał parametryzacji, przyjmowano wartości zalecane przez jego autorów.

3. Model Isinga w odszumianiu dokumentów

Model Isinga, spotykany pod nazwą model Lenza-Isinga, został zaproponowany przez Wilhelma Lenza w 1920 roku [1, 2]. Pod tą nazwą model występował od czasu opublikowania pracy Rudolpha Peierlsa z 1925 roku i odnosił się do przypadku jednowymiarowego [4]. Mimo pierwotnego przeznaczenia jako modelu zjawiska fizycznego model stał się użyteczny również w wielu innych dziedzinach [5, 10]. Model Isinga można określić jako *probabilistyczny model grafowy* (ang. *probabilistic graphical models*). Składa się on z węzłów połączonych ze sobą za pomocą łuków (krawędzi). Jedną z głównych klas tych modeli są grafy nieskierowane, zwane *markowskimi polami losowymi* (ang. Markov Random Field) [1]. Węzły mogą oddziaływać ze sobą w pewnym określonym sąsiedztwie i w obecności zewnętrznego pola. Utworzony graf nieskierowany musi być bez cykli i z co najwyżej jedną krawędzią w parach wierzchołków. Dostosowując model Isinga do potrzeb odszumiania obrazów, oznaczymy przez \mathbf{x} aktualną konfigurację modelu. Zbiór oddziaływań pomiędzy poszczególnymi spinami w sąsiedztwie oznaczymy przez \mathbf{N} . Jeśli dane wartości są ze sobą w sąsiedztwie, to istnieje taka para indeksów m, n , że m i n należą do \mathbf{N} . Model ten można zastosować wprost do modelowania obrazów przez interpretowanie spinów jako pikseli. Co więcej, model pozwala na wygodne zdefiniowanie zależności pomiędzy sąsiadującymi pikselami oraz na określenie związków między elementami obrazu zaobserwowanego i obrazu odszumionego. Ponieważ rozpatrywany obraz jest binarny, przyjęto, że kolor „czarny” jest reprezentowany za pomocą spinu o wartości -1 a kolor „biały” reprezentuje spin $+1$. Rozkład łączny dla grafu określa wzór (1):

$$p(\mathbf{x}) = \frac{1}{Z} \prod_k \psi_k(\mathbf{x}_k), \quad (1)$$

$$Z = \sum_{\mathbf{x}} \prod_k \psi_k(\mathbf{x}_k), \quad (2)$$

$$\psi_k(\mathbf{x}_k) = \exp(-E(\mathbf{x}_k)), \quad (3)$$

gdzie: k oznacza maksymalną klikę w grafie, $\psi_k(\cdot)$ nieujemną funkcję potencjału, natomiast E oznacza funkcję energii.

3.1. Funkcja energii i jej rola w eliminacji szumów

Proponuje się wprowadzenie następującej definicji funkcji energii $E(\mathbf{x}, \mathbf{y})$ w kontekście modelu odszumiania obrazów [10]:

$$E(\mathbf{x}, \mathbf{y}) = h \sum_i x_i + \beta \sum_{(i,j) \in \mathcal{N}} x_i x_j + \eta \sum_i x_i y_i \quad (4)$$

Parametr h w zależności (4) jest składową stałego pola i pozwala na przesunięcia preferencji wyboru wartości spinu. Jest on określany liczbą rzeczywistą. Zakres wartości parametrów η oraz β mieści się w przedziale domkniętym $[0,1]$. Wybór zakresu jest silnie związany z odwzorowaniem danej składowej na zachowanie w odniesieniu do obrazów. Czynniki te reprezentuje miarę stopnia powiązania obrazu odszumionego z obrazem zaobserwowanym. Parametr β pozwala na kontrolę siły wiązań pomiędzy węzłami wchodzącymi w skład danego sąsiedztwa. Mając funkcję energii, można zdefiniować rozkład prawdopodobieństwa $p(\mathbf{x}, \mathbf{y})$ następująco:

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{Z} \exp(-E(\mathbf{x}, \mathbf{y})), \quad (5)$$

$$Z = \sum_{\mathbf{x}} \exp(-E(\mathbf{x})). \quad (6)$$

Celem eliminacji szumów z obrazów jest znalezienie takiego obrazu $\hat{\mathbf{x}}$ na podstawie zaobserwowanych danych \mathbf{y} , który będzie najbardziej zbliżony do oryginału \mathbf{x}^* . W kontekście modelu Isinga problem ten definiujemy jako problem optymalizacji bez ograniczeń, w którym szukamy $\hat{\mathbf{x}}$, maksymalizując rozkład łączny $p(\mathbf{x}, \mathbf{y})$. Okazuje się, że problem ten jest równoważny minimalizacji energii $E(\mathbf{x}, \mathbf{y})$, co można udowodnić na drodze prostych przekształceń algebraicznych.

3.2. Rozszerzenie algorytmu ICM

Do optymalizacji poszukiwania obrazu odszumionego najbardziej zbliżonego do obrazu oryginalnego zastosowano zmodyfikowaną wersję algorytmu *iteracyjnej zmiany warunkowej mody* (ang. Iterated Conditional Modes, ICM) minimalizowania funkcji energii [3]. Konieczność kosztownego obliczeniowo wyznaczania funkcji energii jest dominującym czynnikiem determinującym prędkość uzyskania zgodności obrazów. Ponieważ rozpatrywane są obrazy dokumentów, możliwe jest wykorzystanie ich charakterystycznych właściwości,

tj. dużych obszarów jednolitego koloru, w których piksele obrazu nie zmieniają swojej wartości w bieżącej iteracji i nie trzeba tych wartości wyznaczać. W rozszerzeniu proponuje się zastosowanie pomocniczej macierzy \mathbf{M} , która przechowuje informację o tym, czy dane miejsce jest jednolite względem sąsiedztwa. Aktualizacja maski odbywa się po każdorazowej zmianie stanu modelu \mathbf{x} . Dzięki temu unika się sytuacji przedwczesnej zbieżności spowodowanej zablokowaniem możliwości modyfikacji pikseli obrazu. Maska jest generowana jako splot jądra, które reprezentuje sąsiedztwo, z czarno-białą wersją bieżącego przybliżenia obrazu wynikowego $\tilde{\mathbf{x}}$ (równanie 7)). Przykładowo dla sąsiedztwa typu „+” o rozmiarze 3x3 jądro jest następujące:

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{M} = \mathbf{K} * \tilde{\mathbf{x}} \quad (7)$$

W każdym kroku aktualizowana jest wartość pojedynczego piksela. Możliwe jest także iteracyjne aktualizowanie macierzy maski. Niech N_i oznacza sąsiedztwo wokół punktu i , wliczając w to parę (i,i) . Uaktualnianie wartości m_j dla wszystkich elementów maski, które znajdują się w sąsiedztwie dla i -tego piksela, wartość maski zmienia się zgodnie ze wzorem (7), przy czym:

$$\forall_{(i,j) \in N_i} m'_j = m_j - 2x_i \quad (8)$$

Aktualizacja wartości maski zachodzi jedynie w momencie, gdy zmienia się wartość x_i , co pozwala unikać czasochłonnego procesu jej generowania po każdej aktualizacji wartości piksela.

3.3. Sąsiedztwo

W zależności od zdefiniowania sąsiedztwa zmienia się rozkład oddziaływań pomiędzy pikselami, co prowadzi do bezpośredniego oddziaływania na zmianę energii w zależności od otoczenia punktu. Przyjęto założenie, że każdy rodzaj sąsiedztwa jest kwadratowym obszarem zwanym *oknem* (ang. *window*), dla którego definiowane są wiązania pomiędzy punktem a jego sąsiadami. Rozmiar okna musi być liczbą nieparzystą, a najmniejsza jego wielkość to 3x3. Zdefiniowano trzy podstawowe struktury sąsiedztwa. Sąsiedztwo typu „+” jest jednym z najprostszych sposobów zdefiniowania sąsiedztwa. Do zbioru par sąsiadów zostają dodane wszystkie indeksy znajdujące się na odcinku od punktu z indeksem do granicy wyznaczonej przez rozmiar okna, bez połączeń ukośnych. Drugim rodzajem jest sąsiedztwo typu „x”. Jest to typ pierwszy sąsiedztwa poddany obrotowi o 45°. To pary indeksów, które są położone na przekątnych obszaru wyznaczonego przez okno. Trzecim stosowanym rodzajem jest

sąsiedztwo typu „*”, które powstaje przez złożenie dwóch poprzednich rodzajów sąsiedztwa, czyli typu „+” oraz „x”.

4. Wyniki z badania algorytmów

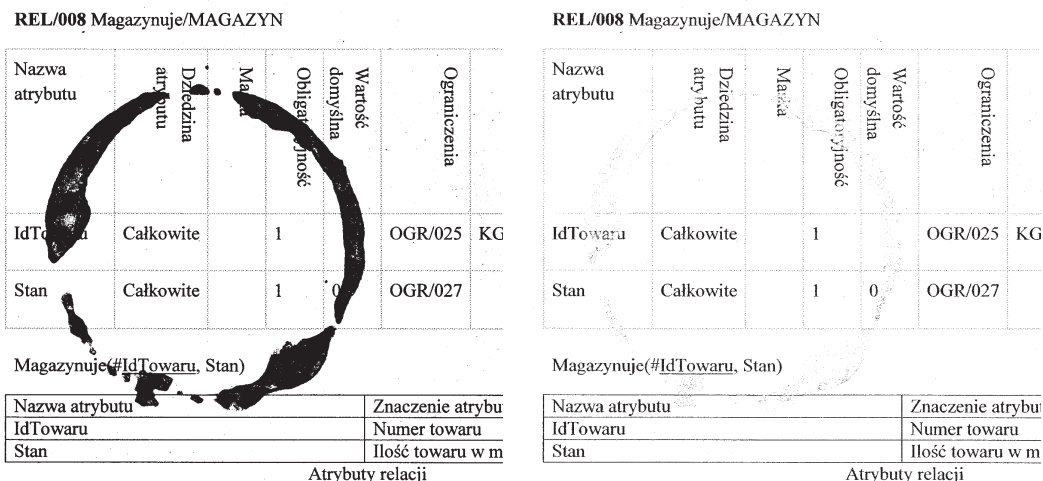
Rezultaty przeprowadzonego badania algorytmów porównano za pomocą trzech miar porównawczych. Pierwszą z nich jest jedna z najczęściej stosowanych miar podobieństwa obrazów, czyli szczytowy stosunek sygnału do szumu (ang. Peak Signal-to-Noise Ratio, PSNR) [14]. Druga z miar opiera się na macierzy błędów (ang. Confusion Matrix) i jest nazywana współczynnikiem kappa (KAPPA). Metoda ta została zaproponowana w 1960 roku przez Jacoba Cohena [6] i jest miarą zgodności pomiędzy dwoma respondentami klasyfikującymi N elementów do C wzajemnie wykluczających się kategorii. Trzecia jest jedną z ostatnio opracowanych metod wyliczania podobieństwa obrazów z wykorzystaniem podobieństwa struktury (ang. Structural Similarity, SSIM) [14]. Jak podają autorzy, metodę tę można potraktować jako próbę rozwiązania problemu, tak aby zwracany współczynnik podobieństwa obrazów odpowiadał klasyfikacji podobieństwa obrazów przez człowieka. Badanie przeprowadzono, poszukując najlepszej kombinacji parametrów spośród trzech rodzajów sąsiedztwa, jego rozmiarów od 3×3 do 9×9 przy stałej wartości $h=0$ i wartościami β i η zawierającymi się w zbiorze $\{0.1, 0.3, 0.7, 1\}$. Zbiorcze wyniki badania algorytmów dla trzech wybranych metryk porównawczych przedstawiono w tabeli 1. Przykładowe dokumenty odszumione zaproponowaną nową metodą znajdują się na rysunkach 1 i 2.

Tabela 1

Zbiorcze wyniki oceny algorytmów odszumiania obrazów dokumentów dla trzech miar porównawczych

MODEL	PSNR		KAPPA		SSIM	
	Średnia	Odchylenie	Średnie	Odchylenie	Średnia	Odchylenie
AVG	14,771	4,703	0,706	0,212	0,798	0,232
BM3D	15,904	4,950	0,734	0,217	0,813	0,226
DWT	9,407	2,750	0,203	0,113	0,659	0,214
ISING	15,966	4,910	0,737	0,216	0,824	0,222
MED	15,551	4,943	0,717	0,219	0,804	0,236
NL	11,915	3,725	0,580	0,191	0,730	0,218
PDE	15,851	4,966	0,733	0,217	0,807	0,229
ROF	13,671	4,356	0,664	0,206	0,772	0,227
SURE	10,154	3,447	0,463	0,178	0,692	0,219
TV	15,861	4,962	0,733	0,217	0,808	0,229
WIE	15,885	4,976	0,730	0,220	0,814	0,233
BM3D	15,904	4,950	0,734	0,217	0,813	0,226
DWT	9,407	2,750	0,203	0,113	0,659	0,214

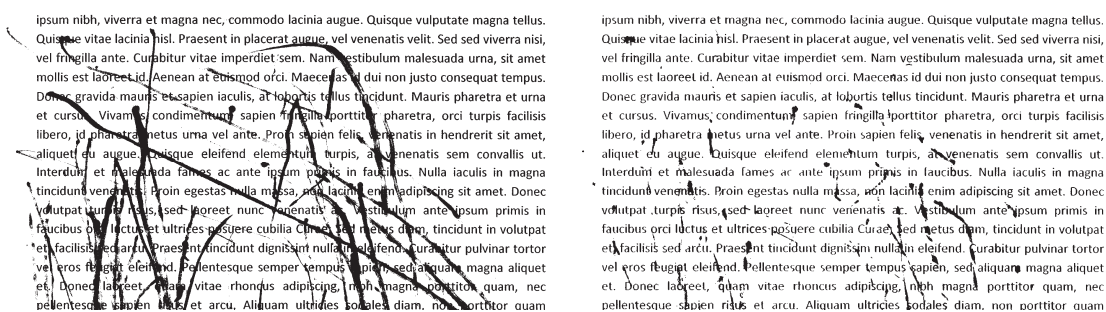
Źródło: opracowanie własne.



Rys. 1. Obraz dokumentu zniszczonego zaplamieniem (po lewej) i obraz odszumiony dokumentu (po prawej)

Fig. 1. Damaged document picture (on the left) and denoising document picture (on the right)

Źródło: opracowanie własne.



Rys. 2. Obraz dokumentu zniszczonego pokreśleniem (po lewej) i obraz odszumiony dokumentu (po prawej)

Fig. 2. Damaged document picture (on the left) and denoising document picture (on the right)

Źródło: opracowanie własne.

Z przeprowadzonych badań wynika, że wykorzystanie modelu Isinga do odszumiania dokumentów daje bardzo dobre rezultaty. Ocenie podlegały najlepsze z otrzymywanych wyników dla każdej pary obrazów testowych, przetwarzanych za pomocą każdego z porównywanych algorytmów. Dla każdej z metryk podobieństwa zaproponowany algorytm oparty na modelu Isinga daje najlepsze rezultaty, zatem może być stosowany do odszumiania obrazów mających zakłócenia fizyczne.

Algorytm może być stosowany jako pomoc przy restaurowaniu i odtwarzaniu typowych uszkodzeń biurowych dokumentów i może być narzędziem przydatnym dla przedsiębiorstw prowadzących procesy przetwarzania i repozytoria dokumentów papierowych. Odszumianie obrazów z fizycznych zagniecia i plam może być wykorzystywane również w bibliotekach i archiwach w procesie ich digitalizacji.

Bibliografia

1. Barber D.: Bayesian Reasoning and Machine Learning. Cambridge University Press, 2012.
2. Bishop C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics, Springer-Verlag, New York, Inc., 2006.
3. Besag J.: On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society* 1986, vol. 48, no. 3, p. 259-302.
4. Brush S.G.: History of the Lenz-Ising Model. *Rev. Mod. Phys.*, American Physical Society 1967, vol. 39, no. 4.
5. Chan S.H., Khoshabeh R., Gibson K.B., Gill P.E., Nguyen T.Q.: An augmented Lagrangian method for total variation video restoration. *IEEE Transactions on Image Processing* 2011, vol. 20, no. 11, p. 3097-3111.
6. Cohen J.: A Coefficient of Agreement for Nominal Scale. *Educational and Psychological Measurement* 1960, vol. 20, no. 1.
7. Dabov K., Foi A., Katkovnik V., Egiazarian K.: Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* 2007, vol. 16, no. 8, p. 2080-2095.
8. Darbon J., Cunha A., Chan T.F., Osher., Jensen G.J.: Fast nonlocal filtering applied to electron cryomicroscopy. *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2008, IEEE ISBI*, p. 1331-1334.
9. Dengwen Z., Wenggang C.: Image denoising with an optimal threshold and neighbouring window. *Pattern Recognition Letters, Elsevier* 2008, vol. 29, no. 11, p. 1694-1697.
10. MacKay D.J.C.: *Information Theory, Inference & Learning Algorithms*. Cambridge University Press, 2002.
11. Mohammed M.S., Saddar P.Y.: Using Discrete Wavelet Transform and Wiener filter for Image De-noising. *Wasit Journal for Science & Medicine* 2009, vol. 2, no. 1, p. 18-30.
12. Rudin L.L., Osher S., Gatemi E.: Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena, Elsevier* 1992, vol. 60, no. 1, p. 259-268.
13. Shannon C.E.: Communication in the presence of noise. *IEEE Proceeding of the IRE* 1949, vol. 37, no. 1, p. 10-21.
14. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 2004, vol. 13, no. 4, p. 600-612.
15. You Y.L., Kaveh M.: Fourth-order partial differential equations for noise removal. *IEEE Transactions on Image Processing* 2000, vol. 9, no. 10, p. 1723-1730.

Abstract

Physical distortions in documents are common. Not many papers are focused on restoring original true image from observed physically damaged documents. Proposed method is focused specific on this task.

Conducted empirical experiments shows that proposed method is superior to other methods in denoising damaged documents. Using proposed ICM extensions it is possible to reduce computation time.

Proposed method could be used as first preprocessing phase in process of restoring old documents in archives, public libraries or in industry. Digitizing old document could help improve their accessibility.