

Image Restoration

Benjamin Steber

11.01.2011

Inhaltsverzeichnis

1	Einleitung	2
2	Grundlagen	2
3	Konstruktion des TV recovery Algorithmus	3
3.1	Fourier Transformation	3
3.2	Optimierungsproblem	4
3.3	Die TV-Norm	4
3.4	Aspekte der Implementierung	5
4	Zusammenfassung	6

1 Einleitung

In Abschnitt III des Buches Level Set Methods and Dynamic Implicit Surfaces von Stanley Osher und Ronald Fedkiw geht es um Bildverarbeitung (Image Processing), also Vorgängen wie Restaurierung (Denoising, Deblurring), Rekonstruktion oder Kompression. Kapitel 11 handelt von der PDE-basierten Restaurierung, bzw der Entwicklung eines Algorithmus zur Restaurierung von verrauschten und unscharfen Bildern durch lösen eines nichtlinearen total-variation Optimierungsproblems.

2 Grundlagen

Warum muss ein Bild aufbereitet werden?

Unschärfe (blurring) entsteht z.B. wenn eine Kamera die wahre Intensität eines Signals nicht richtig auf das Bild abbildet, z.B. verursacht durch Bewegung.

Rauschen (noise) Störgröße, die aufgenommene Intensitäten vom wahren Wert abweichen lässt.

Wie wird das Bild mathematisch dargestellt?

Folgende Erkenntnisse beschränken sich auf die Aufbereitung von 2D Graustufenbildern. Wir beschränken uns also auf den \mathbb{R}^2 .

Die grundlegende und vereinfachende Idee ist es, das Graustufenbild als eine reelle Funktion $u(x, y)$ definiert auf dem Quadrat $\Omega : \{(x, y) | 0 \leq x, y \leq 1\}$ anzusehen, sodass u diskrete Werte zwischen 0 (schwarz) und 255 (weiß) annimmt. Zur Vereinfachung nehmen wir ein Kontinuum an (in Matlab zum Beispiel einfach Werte in $[0, 1]$). Im Folgenden bezeichnet u das Originalbild und u_0 das beobachtete Bild.

Das Modell für Unschärfe kann erzeugt werden durch eine Beeinträchtigung des Originalbilds durch einer Art Mittelung. D.h. z.B. durch die Faltung mit einem Gauß'schen Kern $j(x, y)$:

$$j(x, y) = \frac{1}{4\pi\sigma} e^{-(x^2+y^2)/4\sigma}$$

mit den Eigenschaften:

$$\begin{aligned} j(x, y) &\geq 0 \\ j(x, y) &\rightarrow 0 \text{ wenn } (x^2 + y^2)^{1/2} \rightarrow \infty \\ \int_{\mathbb{R}^2} j(x, y) dx dy &= 1 \end{aligned}$$

und man erhält:¹

$$u_0(x, y) = (j * u)(x, y) = \int_{\Omega} u(s, r) j(x - s, y - r) ds dr \quad (2.1)$$

Für das Rauschen setzen wir ein additives weißes gauß'sches Rauschen n an, wobei die n_i von n zum Pixel i unabhängige normalverteilte Zufallsvariablen sind mit Erwartungswert 0 und Varianz σ^2 . Für das gesamte Modell des unscharfen und verrauschten Bildes erhält man also:

$$j * u + n = u_0 \quad (2.2)$$

¹Kapitel 11.1 beschreibt dass u_0 durch Lösung der Wärmeleitungsgleichung gewonnen werden kann.

Definition (signal-to-noise ratio):

$$SNR = \frac{\|u - \bar{u}\|_{\mathcal{L}^2}}{\sigma}$$

wobei \bar{u} der Quotient aus Standardabweichung des Signals und der Standardabweichung des noise ist.

Und eine angewandtere Definition:

$$SNR = -20 \log_{10} \left(\frac{\|u - u_0\|}{\|u\|} \right)$$

Das SNR ist ein Maß für den Fehler zwischen dem Originalbild und dem verfälschten Bild. Je höher das SNR umso besser die Qualität des Bildes. Dies dient nur zum Test des Algorithmus, da im Anwendungsfall das Original unbekannt ist.

3 Konstruktion des TV recovery Algorithmus

3.1 Fourier Transformation

Ausgehend von der Modellannahme

$$j * u + n = u_0$$

und dem beobachteten Bild u_0 soll das ursprüngliche Bild wieder gewonnen werden. Eine übliche Herangehensweise ist es die Fouriertransformation von (2.1) zu nehmen. Der *Faltungssatz* besagt, dass die Fourier-Transformierte einer Faltung gleich dem Produkt der Fourier-Transformierten der beiden Funktionen entspricht. Also wird die Faltung zur einfachen Multiplikation.

$$\hat{u}_0(k, l) = \hat{j}(k, l) \hat{u}(k, l) \quad (3.1)$$

Um das Modell (2.2) zu lösen, nimmt man die Fourier Transformation und kommt zu:

$$\hat{j}(k, l) \hat{u}(k, l) + \hat{n}(k, l) = \hat{u}_0(k, l) \quad (3.2)$$

Um das Bild $u(x, y)$ zu erhalten, könnte man nun nach u auflösen und rücktransformieren.

$$\hat{u}(k, l) = \frac{\hat{u}_0(k, l) - \hat{n}(k, l)}{\hat{j}(k, l)} \quad (3.3)$$

Das ist aber grundsätzlich ein sehr schlecht gestelltes Problem. Zum einen existiert keine eindeutige analytische Lösung und da $\hat{j}(k, l) \rightarrow 0$ schnell wenn $(k^2 + l^2)^{1/2} \rightarrow \infty$ werden hohe Frequenzen in u_0 stark verstärkt, denn \hat{j} fällt meist gerade zu hohen Frequenzen hin gegen 0, das Rauschen hat aber auch dort einen Anteil der dann durch $1/\hat{j}$ verstärkt wird. Zudem ist das Bild u_0 stückweise glatt mit Sprüngen sodass bei der Fouriermethode das Gibbsche Phänomen auftritt (siehe Abb. 1).

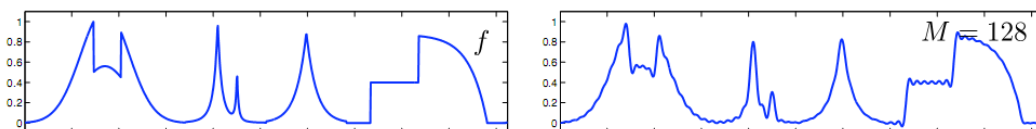


Abbildung 1: Ringing bei Fourier Transformation

3.2 Optimierungsproblem

Um die oben genannten Probleme zu umgehen, schlagen die Autoren nun eine Regularisierung des Problems vor, indem man Vorabkenntnisse über das Bild (Glattheit, Existenz von Kanten), dem Rauschen und der Unschärfe in ein Funktional R packt, das die Qualität eines Bildes u misst, sodass kleine Werte von $R(u)$ einem guten Bild angehören. Also ein Optimierungsproblem mit Nebenbedingungen:

$$\begin{aligned} \min_u R(u) \\ \text{mit } \|j * u - u_0\|_{\mathcal{L}^2}^2 = |\Omega|\sigma^2 \end{aligned} \quad (3.4)$$

Die Bedingung erklärt sich folgendermaßen:

Nach dem Modell (2.2) ist $n = u_0 - j * u$ und der Erwartungswert³ $E(\int_{\Omega} n^2 d\Omega) = |\Omega|\sigma^2$ daraus folgt $\|j * u - u_0\|_{\mathcal{L}^2}^2 = \int_{\Omega} (j * u - u_0)^2 d\Omega \approx |\Omega|\sigma^2$.

Angegebene Beispiele für regularisierende Funktionale sind zum Beispiel $R(u) = \|\Delta u\|_{\mathcal{L}^2}, \|\nabla u\|_{\mathcal{L}^2}$.

3.3 Die TV-Norm

Der Nachteil der oben genannten Funktionale ist, dass sie keine Unstetigkeiten in der Lösung erlauben. D.h. dort wo klare Kanten erwünscht sind werden diese nicht optimal wiederhergestellt.

Die "total variation" (TV) einer reellen Funktion f auf einem Intervall $[a, b]$ ist definiert als:

$$TV(f) = \sup_P \sum_{i=0}^{n_P-1} |f(x_{i+1}) - f(x_i)|$$

wobei das Supremum über die Menge aller Partitionen $\mathcal{P} = \{P = \{x_0, \dots, x_{n_P}\} | P \text{ Partition von } [a, b]\}$ läuft.⁴

Für höhere Dimensionen gibt es eine entsprechende Definition, für die betrachteten Bilder interessiert aber nur die erste und zweite und es kann gezeigt werden, dass für unser Bild u die TV-Norm folgendermaßen definiert werden kann:

$$TV(u) = \int_{\Omega} |\nabla u| d\Omega = \int_{\Omega} \sqrt{u_x^2 + u_y^2} d\Omega$$

Die TV-Norm bestraft jetzt nicht die Unstetigkeitsstellen in u und ermöglicht es so die Kanten des Originalbildes wiederherzustellen.

Damit kann (3.4) folgendermaßen formuliert werden:

$$\begin{aligned} \min_u \int_{\Omega} |\nabla u| d\Omega \\ \text{mit } \frac{1}{2} \left(\int_{\Omega} (j * u - u_0)^2 d\Omega - |\Omega|\sigma^2 \right) = 0 \end{aligned} \quad (3.5)$$

Die Lagrangefunktion davon ist:

$$\int_{\Omega} |\nabla u| d\Omega + \frac{\lambda}{2} \left(\int_{\Omega} (j * u - u_0)^2 d\Omega - |\Omega|\sigma^2 \right) \quad (3.6)$$

Wie auch im Kapitel über Snakes, Active Contours and Segmantation sucht man nun lokale Extremalstellen der Lagrangefunktion durch Nullsetzen ihres Gradienten:

³ $|\Omega|$ könnte man hier auch weglassen weil 1.

⁴Definition von Wikipedia http://en.wikipedia.org/wiki/Total_variation

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(j * (j * u - u_0)), \quad (3.7)$$

$$0 = \frac{1}{2} \left(\int_{\Omega} (j * u - u_0)^2 d\Omega - |\Omega| \sigma^2 \right) \quad (3.8)$$

(3.8) ist wieder die Nebenbedingung.

Um solche Optimierungsprobleme mit Nebenbedingungen wie (3.4) zu lösen, gibt es Techniken, die die Löser für das unrestringierte Problem benutzen, deren Euler-Lagrange Gleichung (3.7) mit festem λ sind. Deshalb gehen wir jetzt davon aus, dass wir λ kennen. Für $\nu = \frac{1}{\lambda}$ erhält man das äquivalente unrestringierte Problem durch

$$\min_u \int_{\Omega} \left(\nu |\nabla u| + \frac{1}{2} (j * u - u_0)^2 \right) d\Omega \quad (3.9)$$

und erhalten auf dem gleichen Weg wie oben daraus die Euler-Lagrange Gleichung:

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda j * (j * u - u_0), \quad (3.10)$$

was die Autoren Marquina und Osher “*nonlinear deconvolution model*” nennen.⁵ Das Gradientenverfahren ist die Standardmethode dies zu lösen:

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda j * (j * u - u_0) \quad (3.11)$$

Und man hat damit einen TV-basierten denoising/deblurring Algorithmus gegeben.

3.4 Aspekte der Implementierung

Das so entwickelte Verfahren sollte nun numerisch effektiv implementiert werden. Gleichung (3.11) kann wie folgt interpretiert werden. Der nicht λ Term bewegt jede Niveaulinie von u in Normalenrichtung mit Geschwindigkeit gleich ihrer mittleren Krümmung geteilt durch die Länge des Gradienten, während der λ -Term die Nebenbedingung enthält um einen nicht trivialen stationären Zustand zu erhalten und sich keine Verzerrung des eigentlichen Bildes einstellt. Der erste Term entfernt die stark gekrümmten Noiseanteile und filtert daraus die stückweise glatten Funktionen.

Wollen wir nur im Fall $\lambda = 0$ (3.11) approximieren können wir die Gleichung so schreiben und müssen aufpassen an Stellen wo $\nabla u = 0$ ist:

$$u_t = \frac{\partial}{\partial x} \left(\frac{u_x}{\sqrt{u_x^2 + u_y^2 + \delta}} \right) + \frac{\partial}{\partial y} \left(\frac{u_y}{\sqrt{u_x^2 + u_y^2 + \delta}} \right) \quad (3.12)$$

wobei $\delta > 0$ klein gewählt ist.

Ein explizites Verfahren sollte nun intuitiv eingeschränkt sein durch

$$\frac{\Delta t}{(\Delta x)^2} \leq c \sqrt{u_x^2 + u_y^2 + \delta} \quad (3.13)$$

Das führt aber zu der Annahme, dass dieses Verfahren sehr langsam in den stationären Zustand konvergiert, gerade wenn das Bild stark verrauscht und unscharf ist. Außerdem führen Fedkiw und Osher noch weiter Probleme an die auftauchen.

⁵entspricht Gleichung (11.23) im Buch

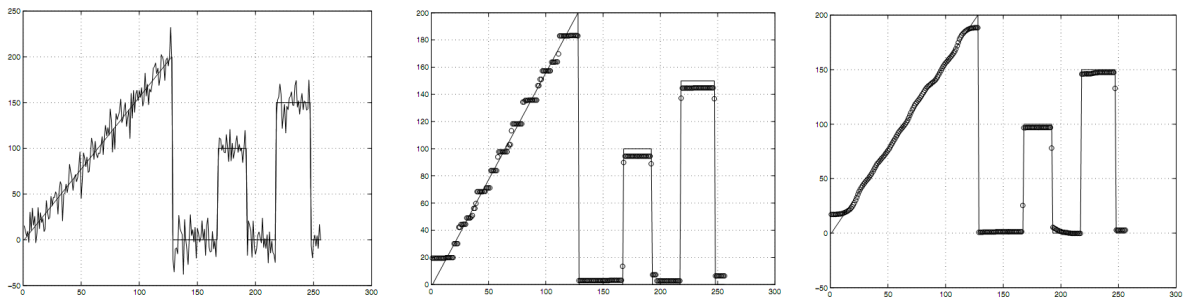
Deshalb schlagen Marquina und Osher noch eine Veränderung vor. Sie multiplizieren die rechte Seite mit $|\nabla u|$

$$u_t = |\nabla u| \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda |\nabla u| j * (j * u - u_0) \quad (3.14)$$

was mehrere Vorteile birgt. Als Zeitschritt Restriktion kann jetzt $\Delta t / (\Delta x)^2 \leq c$ für ein $c > 0$ gewählt werden, sodass einfache explizite Verfahren benutzt werden können. Für den Konvektionsterm kann ENO oder WENO und für den Diffusionsterm kann central differencing benutzt werden. Was wohl zur Folge hat, dass die bei anderen TV-basierten Restaurierungsalgorithmen auftauchende “staircase Effekte” kein Problem darstellen (s. Abb. 2(b))

Um nun mithilfe dieses Verfahrens ein Bild zu rekonstruieren, schlagen Osher und Fedkiw vor, im Falle eines nur verrauschten Bildes, dieses als Startbild zu verwenden, hingegen im deblurring Fall zuerst eine lineare Dekonvolution des beobachteten Bildes als erste Näherung zu nehmen.

Abbildung 2



(a) 1D Original Signal und verrauschtes Signal SNR ≈ 3 (b) TV Rekonstruktion Verfahren (3.11) mit staircase Effekten (c) TV Rekonstruktion verbessertes Verfahren (3.14)

4 Zusammenfassung

“Das ist ein weites Feld...”⁶ Insgesamt ist das Feld der Denoising und Deblurring Techniken ein weites und der hier vorgestellte Algorithmus ist nur ein Ansatz unter vielen. Die vorgestellte TV Methode hat aber, im Gegensatz zu vielen anderen L^2 Ansätzen den Vorteil, dass Noise erfolgreich reduziert, dabei aber die für das Bild relevanten Kanten erhalten bleiben. Aber auch die L^2 Ansätze sind schon soweit verbessert und getrimmt worden, dass sie gute Ergebnisse liefern.

⁶Theodor Fontane, Effi Briest