

Effiziente Bildverarbeitung mit primal-dualen Algorithmen

Efficient Image Processing using Primal-Dual Algorithms

Thomas Pock



Thomas Pock ist Assistenzprofessor am Institut für Maschinelles Sehen und Darstellen. Seine Forschungsinteressen liegen im Bereich der Entwicklung von effizienten Optimierungsverfahren für die Bildverarbeitung und das maschinelle Sehen. Seit 2009 ist er Vorsitzender der Österreichischen Arbeitsgemeinschaft für Mustererkennung (OAGM).

Thomas Pock is assistant professor at the Institute for Computer Graphics and Vision. His research interests are in the field of the development of efficient optimization methods for image processing and computer vision. Since 2009, he has been chairman of the Austrian Association for Pattern Recognition (OAGM).

Am Institut für Maschinelles Sehen und Darstellen wird an hocheffizienten primal-dualen Algorithmen zur Lösung von konvexen Optimierungsproblemen in der Bildverarbeitung geforscht. Diese spezialisierten Algorithmen kommen in zahlreichen praktischen Problemen wie Fahrerassistenzsystemen, der dreidimensionalen Gebäuderekonstruktion sowie in der Aufbereitung von medizinischen Bildern zur Anwendung.

Durch die rasante Entwicklung von Bildaufnahme-systemen wie hochauflösenden Digitalkameras, medizinischen Scannern, aber auch Smartphones ist die Bildverarbeitung mittlerweile ein zentrales Thema in Forschung und Industrie geworden. Fast jeder Mensch hat eine Videokamera in Form eines Smartphones in der Tasche. Ultrahochoauflösende Digitalkameras erlauben, ganze Städte aus der Luft aufzunehmen und mithilfe der Methoden des maschinellen Sehens dreidimensional zu rekonstruieren. Auch in der Verkehrssicherheit spielt das maschinelle Sehen zunehmend eine größere Rolle. So wird die Daimler S-Klasse der neuesten Generation bereits mit einem Fahrerassistenzsystem ausgestattet sein, welches überwiegend auf Methoden des maschinellen Sehens beruht (vgl. Abb. 1).

Die Methode des maschinellen Sehens führt zu sogenannten inversen Problemen, das bedeutet, dass man aus den beobachteten Bildern, die zuvor durch eine Projektion der dreidimensionalen Welt auf die zweidimensionale Bildebene entstanden sind, wieder auf Eigenschaften der dreidimensionalen Welt rückschließen möchte. Da diese Probleme im Allgemeinen keine eindeutige Lösung haben, muss ein gewisser Grad an Vorwissen eingebracht werden. Als großes Vorbild fungiert hier das menschliche visuelle System, welches selbst unter widrigsten Umständen ist, wichtige Eigenschaften der Welt zu

The Institute for Computer Graphics and Vision conducts research on highly efficient primal-dual algorithms for convex optimization. These highly specialized algorithms find their applications in numerous practical problems, such as driver assistance systems, three-dimensional reconstructions of buildings and the processing of medical images.

Brought about by the rapid development of imaging hardware, such as high resolution digital cameras, medical scanners and smartphones, image processing has become a central topic in research and industry. These days, almost everyone has a video camera connected to a high-performance computer in form of a smartphone in their pocket. Moreover, ultra-high resolution digital cameras nowadays allow aerial pictures of whole cities to be made, and computer vision algorithms can be used to compute the resulting three-dimensional reconstructions. Computer vision also plays an increasingly important role in road safety. For example, the latest generation of Daimler's S-class will be equipped with a driver-assistance system that mainly relies on computer vision algorithms (Fig. 1).

Problems such as these are often encountered in computer vision and are referred to as inverse problems, since the task is to determine information about the three dimensional world based on the observed images that have been generated prior to this by projecting the three dimensional world onto the image plane. In general, such problems do not have a unique solution, and hence one has to introduce some degree of prior knowledge. The human visual system acts as a great model in that it is able to determine important information about the world even under very adverse conditions (e.g. driving a car in rain and darkness). In order to impose such prior knowledge onto the problem formulation, a common



erkennen (z. B. Autofahren bei Regen und Dunkelheit). Um solches Vorwissen effizient in die Problemformulierung einzubringen, stellt man diese Aufgabe sehr häufig als mathematisches Optimierungsproblem dar. Dabei betrachtet man ein sogenanntes Energiefunktional, welches bewertet, wie sehr eine Lösung dem Vorwissen entspricht und wie gut die beobachteten Daten (Bilder) erklärt werden. Ziel ist es, jene Lösung zu finden, welche das Energiefunktional minimiert, also den besten Kompromiss zwischen Vorwissen und Datentreue liefert.

Optimierungsprobleme wie jene im Bereich des maschinellen Sehens gelten als unlösbar, weil der Raum an möglichen Lösungen zu groß ist, um effizient die beste Lösung zu finden. Eine Ausnahme bildet die spezielle Klasse von sogenannten konvexen Optimierungsproblemen, welche die Eigenschaft hat, dass ein Gradientenabstieg garantiert zu einer Lösung mit dem geringstmöglichen Wert des Energiefunktionals führt. Sinngemäß erlauben solche Funktionale, die global beste Lösung eines Problems zu berechnen. Die Kunst besteht darin, ein bestimmtes Problem als konvexes Optimierungsproblem darzustellen. Konvexe Optimierungsprobleme haben zusätzlich noch die Eigenschaft, dass neben dem „primalen“ Problem ein sogenanntes „duales“ Problem existiert, dessen optimale Lösung denselben Funktionswert wie das primale Problem aufweist. Die zusätzliche Betrachtung des dualen Problems erlaubt es zudem, Nichtdifferenzierbarkeiten des primalen Problems in einfacher zu handhabende Nebenbedingungen umzuwandeln.

Am Institut für Maschinelles Sehen und Darstellen wurde in Zusammenarbeit mit Antonin Chambolle von der École Polytechnique in Paris-Palaiseau ein spezielles gradientenbasiertes Optimierungsverfahren¹ entwickelt, welches gleichzeitig sowohl das primale als auch das duale Problem löst. Dadurch kann man auf elegante Weise

approach is to pose the problem as a mathematical optimization problem. The idea is to consider an energy functional whose functional value expresses, how well a solution is consistent with the prior knowledge and how well the observed data (images) can be explained. The goal is then to find the solution that minimizes the energy functional, thus providing the best compromise between prior knowledge and data fidelity.

Optimization problems like these in computer vision are generally considered as being unsolvable since the space of possible solutions is too big to allow for an efficient identification of the best solution. One exception is the special class of convex optimization problems which, roughly speaking, have the property that a gradual descent always leads to a solution with the lowest possible value of the energy functional. In turn, such functionals allow the best possible solution to a problem to be found. The difficulty is of course to formulate the problem as a convex optimization problem. Convex optimization problems have the additional property that, additional to the actual “primary” problem, there exists a so-called “dual” problem whose optimal solution has the same function value as the primary problem. On top of this, considering this as a dual problem allows non-differentiability of the primary problem to be converted into constraints that are easier to handle.

In co-operation with Antonin Chambolle from the École Polytechnique in Paris-Palaiseau, the Institute for Computer Graphics and Vision has developed a specialized gradient-based optimization method¹ which jointly optimizes the primary as well as the dual problem. This allows, in a smart way, even non-differentiable functionals to be minimized, such as the robust L1 norm. In addition, the method allows convex constraints of various forms to be easily handled. Remarkably, it could be shown that for certain problem classes, the method is optimal in the sense that there

Abb. 1: Bewegungsschätzung aus Bildsequenzen (Kooperation mit Daimler AG). Die Abbildung zeigt links ein Eingabebild und rechts das geschätzte Bewegungsfeld, welches zur besseren Visualisierung farbkodiert wurde (siehe Farbcode). Durch den entwickelten Algorithmus kann die Bewegungsschätzung in Echtzeit erfolgen und somit in Fahrerassistenzsystemen zur Gefahrenerkennung eingesetzt werden² (siehe auch ► www.6d-vision.de).

Fig. 1: Motion estimation in image sequences (co-operation with Daimler AG). On the left-hand side, the figure shows one input image and, on the right-hand side, the estimated motion field which has been color coded for a better visualization (see color code). The developed algorithm allows the motion estimation to be performed in real time and thus to be applied in a driver assistance system to detect dangerous situations² (see also ► www.6d-vision.de).



© Microsoft Photogrammetry

Abb. 2: Stereo aus multiplen Ansichten (Kooperation mit Microsoft Photogrammetry). Die Abbildung zeigt oben die Tiefenkarte (je heller der Grauwert, desto tiefer) und unten eine Visualisierung des finalen 3-D-Modells. Der entwickelte primal-duale Algorithmus erlaubt eine effiziente und global optimale Lösung von schwierigen Stereoproblemen wie der hier gezeigten 3-D-Rekonstruktion des Grazer Kunsthauses³.

Fig. 2: Multi-view stereo (co-operation with Microsoft Photogrammetry). Above, the figure shows the depth map (lighter values indicate more depth), Below, a visualization of the final 3D model. The developed algorithm allows the efficient and globally optimal solution of difficult stereo problems such as a 3D reconstruction of the Grazer Kunsthaus³.

Literatur/References:

¹A first-order primal-dual algorithm for convex problems with applications to imaging, A. Chambolle, T. Pock, *Journal of Mathematical Imaging and Vision* 40 (1), 120–145.

²An improved algorithm for TV-L1 optical flow, A. Wedel, T. Pock, C. Zach, H. Bischof, D. Cremers, *Statistical and Geometrical Approaches to Visual Motion Analysis*, 23–45.

³TGV-Fusion, T. Pock, L. Zebedin, H. Bischof, *Advances in Distributed Agent-Based Retrieval Tools*, 245–258.

selbst nichtdifferenzierbare Funktionale wie die robuste L1-Norm minimieren. Zusätzlich erlaubt das Verfahren, konvexe Nebenbedingungen verschiedener Art sehr einfach zu handhaben. Bemerkenswerterweise konnte gezeigt werden, dass das Verfahren für spezielle Problemklassen optimal ist, das bedeutet, dass es kein anderes gradientenbasiertes Verfahren gibt, welches in derselben Klasse von Problemen effizienter ist. Ein weiterer Vorteil ist die effiziente Parallelisierung. Eine attraktive Plattform sind moderne, programmierbare Grafikkarten, die heutzutage in einer Steckkarte dieselbe Rechenleistung bieten wie hallenfüllende Supercomputer vor zehn Jahren.

is no other gradient-based method that is more efficient in the same class of problems. An additional advantage of the algorithm is that it can be efficiently parallelized. An attractive platform is given by state-of-the-art graphics cards, which, nowadays, provide the equivalent processing power offered by a room-filling super-computer 10 years ago on a small card. The developed primal-dual algorithm has numerous applications in image processing and computer vision. Figures 1 and 2 illustrate two applications. In both examples, the proposed primal-dual algorithm was successfully applied to solve the associated convex optimization problems. ■



Der entwickelte primal-duale Algorithmus hat zahlreiche Anwendungen in der Bildverarbeitung und dem maschinellen Sehen. Abbildungen 1 und 2 illustrieren zwei Anwendungsgebiete. In beiden Beispielen wurde der vorgestellte Algorithmus erfolgreich eingesetzt, um die damit verbundenen konvexen Optimierungsprobleme zu lösen. ■