

# Postproduktion für Kamera-Arrays

*Siegfried Foessel, Wolfgang Hueller, Peter Schäfer, Frederik Zilly*

*Fraunhofer IIS, Abteilung Bewegbildtechnologien  
Erlangen, Deutschland*

fsl@iis.fraunhofer.de

## Zusammenfassung

Jedes Objekt im Raum emittiert und/oder reflektiert Licht. Alle so im Raum befindlichen Lichtstrahlen zusammen bilden ein Lichtfeld, wobei jeder Lichtstrahl mathematisch durch Richtung und Intensität beschrieben werden kann. Bei herkömmlichen Kameras wird eine Vielzahl von Lichtstrahlen gebündelt, sodass nur bestimmte Tiefenschärfen und Abbildungsrichtungen erzeugt werden können und bereits bei der Aufnahme festgelegt werden. Bei geeigneter separater Aufnahme der Lichtstrahlen sind jedoch nachträgliche Schritte und Spezialeffekte in der Postproduktion möglich. Separierende Aufnahmesysteme sind plenoptische Kameras oder Kamera-Arrays. Am Fraunhofer-Institut für Integrierte Schaltungen IIS wird derzeit an einem neuen Lichtfelderfassungs- und Datenverarbeitungssystem auf Basis eines dünn besetzten Kamera-Arrays geforscht, d. h. nur wenige in einem 2D-Array angeordnete hochauflösende Kameras erfassen die Szene. Der Vortrag erläutert die algorithmischen Schritte zur Nachbearbeitung der Szene – wie Rektifizierung, Tiefenschätzung, Disparitätsschätzung usw. – und demonstriert die möglichen Spezialeffekte. Im Rahmen einer Masterarbeit wurde ein Stop-Motion-Film gedreht, bei dem visuelle Effekte wie Ansichtsänderung und Schärfeverlagerung in der Postproduktion durchgeführt wurden, um diese Möglichkeiten zu testen.

## 1 Einleitung

Herkömmliche Filmkameras versuchen die Sicht des menschlichen Auges nachzubilden. Die Kamera beobachtet die Szene von einer bestimmten Position aus und fokussiert auf ein bestimmtes Objekt, so wie es der Mensch bei aufmerksamer Beobachtung auch tun würde. Kameraleute in der Filmproduktion versuchen den Blickwinkel des Menschen nachzubilden und nach ihren Vorstellungen zu führen. Diese kreative Art des Geschichtenerzählens erfordert eine sehr genaue Arbeit am Set, da die Anzahl der Aufnahmeversuche nicht beliebig erhöht werden kann. Die

che nicht beliebig erhöht werden kann. Die optimale Kameraeinstellung während der Aufnahme muss zeitlich mit der bestmöglichen Darstellungsleistung des Schauspielers zusammenpassen.

Die Fragen hierbei sind: Kann der Prozess der Aufnahme am Set zuverlässiger gestaltet werden, sodass die beste Performance des Schauspielers auf jeden Fall optimal erfasst wird? Können die kreativen Möglichkeiten des Kameramannes erweitert werden, trotz des heutigen Zeitdrucks am Set? Eine mögliche Antwort könnte der Einsatz der Lichtfeldtechnologie sein. Ein Lichtfeld ist das Feld aller Lichtstrahlen im Raum, das – wenn es erfassbar wäre – nahezu unbegrenzte gestalterische Möglichkeiten bietet. Sollten alle Lichtstrahlen einzeln erfassbar sein, könnten zum Beispiel der Blickwinkel und die Tiefenschärfe nachträglich festgelegt und geändert werden. Leider ist eine vollständige Erfassung eines Lichtfeldes nicht möglich, aber es können Teile des Lichtfeldes abgetastet werden und damit neue Freiheitsgrade in der Gestaltung geschaffen werden.

In dieser Präsentation wird ein neues Lichtfelderfassungs- und -verarbeitungssystem vorgestellt, welches den konventionellen Kameraansatz mit festgelegter Betrachtungsperspektive verlässt und zu einem flexibleren Ansatz der Lichtfelderfassung einer Szene wechselt. Die Flexibilität in der Postproduktion wird dadurch – wenn auch nur teilweise – erhöht.

## **2 Der Bilderfassungsprozess**

Das Ergebnis einer Filmproduktion ist ein Film, der eine Geschichte aus einer bestimmten Perspektive erzählt. Die Zusammensetzung und die Blickwinkel der Bilder sind dabei entscheidend, um die Aussagen des Films gut vermitteln zu können. Herkömmliche Kameras bündeln eine Auswahl von Lichtstrahlen einer Szene, um das Einzelbild zu formen. In den frühen Tagen der Fotografie wurden Lochkameras verwendet. Die Größe des Loches definiert, wie viele der Lichtstrahlen in die Kamera einfallen. Je kleiner das Loch, desto schärfer wird das Kameraabbild auf der fotoempfindlichen Schicht, da weniger Lichtstrahlen aus unterschiedlichen Richtungen gesammelt werden und so das Abbild nicht „verschmieren“ können. Allerdings führen kleinere Löcher auch zu einer geringeren Kameraempfindlichkeit; die Aufnahmezeit muss verlängert werden und es machen sich irgendwann Beugungseffekte bemerkbar. Eine Lösung für dieses Problem war die Einführung von Objektiven vor der Kamera. Objektive fokussieren Lichtstrahlen auf einer Brennebene. Dies gilt allerdings nur für Lichtstrahlen von Objekten in einer bestimmten Entfernung. Objekte in anderem Abstand zur Kamera wer-

den unscharf abgebildet. Das Maß der Unschärfe wird durch den Unschärfekreis bestimmt.

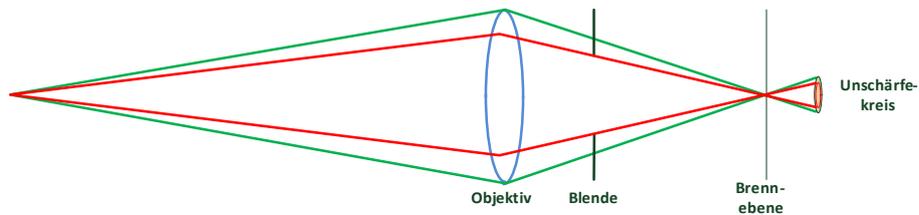


Abb. 1 Abbildung eines Objektpunktes auf den Sensor

Kleinere Blenden führen zu einer größeren Schärfentiefe, da der Unschärfekreis kleiner wird. Größere Objektive und offenere Blenden können mehr Lichtstrahlen einfangen und abbilden, der Unschärfekreis wird größer, die Schärfentiefe wird reduziert. Mit Lichtfeld-Kameratechnologien können Lichtstrahlen individueller erfasst werden, sodass neben der Intensität auch teilweise Richtungsinformationen der Lichtstrahlen erhalten bleiben.

### 3 Lichtfelderfassung

Lichtfelder sind elektromagnetische Felder. In jedem Raumpunkt  $(x, y, z)$  existieren Lichtstrahlen der Intensität  $L(x, y, z, \theta, \phi)$  in unterschiedlicher Raumrichtung  $(\theta, \phi)$ .

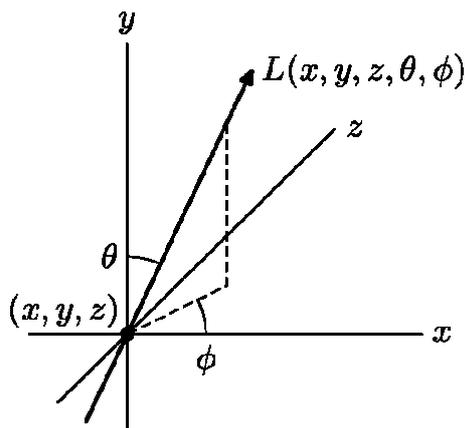


Abb. 2 Definition des Lichtfeldes

(Quelle: Wikipedia, Stichwort „Lichtfeld“)<sup>1</sup>

<sup>1</sup> <http://de.wikipedia.org/wiki/Lichtfeld> <2013-12-08>

Eine vollständige Erfassung des Lichtfeldes ist zwar nicht möglich, da die Lichtstrahlen durch Messung aufgehalten werden, jedoch eine teilweise Abtastung. In diesem Artikel werden zwei Arten der Lichtfelderfassung diskutiert: eine basierend auf Lichtfeldkameras mit Mikrolinsen, die andere basierend auf Kamera-Arrays.

### 3.1 Lichtfelderfassung basierend auf Kameras mit Mikrolinsen

Der erste Ansatz basiert auf der Lichtfelderfassung mithilfe von Kameras mit Mikrolinsen. Die Mikrolinsen werden dabei direkt vor dem Sensor in den Strahlengang eingeführt. Das Hauptobjektiv erzeugt ein Zwischenluftbild in einer Brennebene zwischen Objektiv und Sensor. Das Zwischenluftbild wird mit den Mikrolinsen mehrfach auf den Sensor abgebildet. Dadurch erhält man eine teilweise Auftrennung der Richtungsinformation der einzelnen Lichtstrahlen.

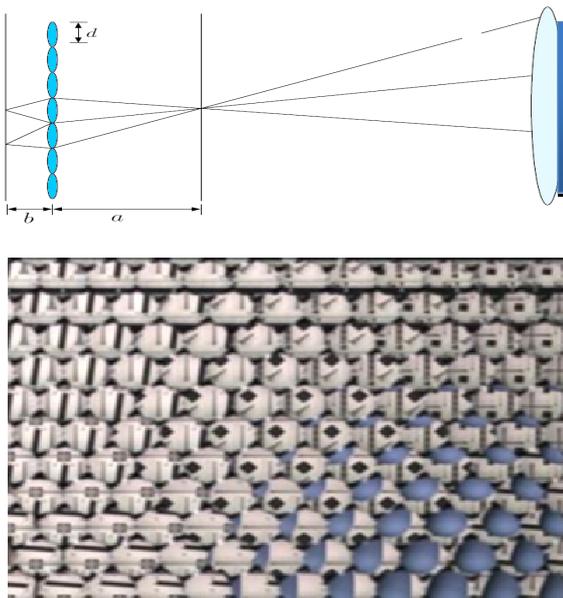


Abb. 3 Lichtfelderfassung mit Kamera und Mikrolinsen-Array

Dabei ergaben sich jedoch einige Nachteile:

- Der Sensor liefert ein Bild in mehrfacher Abtastung. Dies ist notwendig, um hinterher die Lichtfeldverarbeitung durchführen zu können. Allerdings bedeutet dies, dass der Sensor das Szenenabbild in redundanter Weise mehrfach aufnimmt. Im speziellen Fall wurde die Szene von einem 5-MPixel-Sensor fünffach in jeder Richtung mehrfach abgebildet.

Aus einem 5-MPixel-Bild entsteht daher am Ende nur noch ein finales Bild mit VGA-Auflösung.

- Das optische System muss sehr exakt montiert und kalibriert werden, insbesondere durch die verkleinerten Abbildungen zwischen der Hauptoptik und dem Sensor. Außerdem wird über das Mikrolinsen-Array die optische Abbildung einmal unwiderruflich festgelegt.
- Die Verschiebung der Blickpunkts-Perspektive bei der Generierung eines virtuellen Bildes aus dem Lichtfeld ist begrenzt auf den Blickwinkel der Kamera und damit des Sensors. Damit sind nur sehr kleine Winkeländerungen in der Perspektive möglich.

### 3.2 Lichtfelderfassung mit Kamera-Arrays

Ein weiterer Ansatz ist die Verwendung eines Kamera-Arrays. Der Ansatz mit Mikrolinsen wird sozusagen makroskopisch erweitert. Im speziellen Beispiel wurde ein Kamerasystem auf der Basis von netzwerkbasierten Kameras entwickelt. Dies reduziert den Verdrahtungsaufwand.

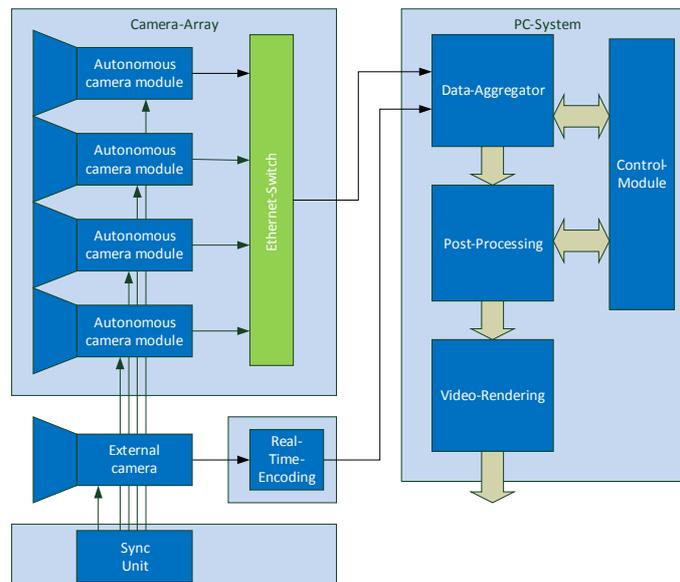


Abb. 4 Systemkomponenten eines Lichtfeldproduktionssystems mit Kamera-Array (Quelle: Fraunhofer IIS)

Die Kameras werden über einen Ethernet-Switch angesteuert und deren Daten über die Ethernet-Schnittstelle abgerufen. Ein Zentralrechner dient der Steuerung und der weiteren Verarbeitung. Auch eine Option zur Integration von Dritthersteller-Kameras wurde realisiert.



Abb. 5 Kamera-Array-System

Die Kameras sind in einem regelmäßigen Raster angeordnet (zum Beispiel  $2 \times 2$  oder  $4 \times 4$  Kameras). Dies erleichtert die Verrechnung der Kameradaten, da die Positionen und Anordnung der Kameras vorher bekannt sind.

## 4 Lichtfeldverarbeitung

Die Kameras innerhalb des Arrays tasten das Lichtfeld nur an bestimmten Punkten ab (dünn besetztes Lichtfeldmodell). Aus dem dünn besetzten Lichtfeld wird ein dichtes Lichtfeld errechnet. Dies erlaubt später eine einfachere Verrechnung der Lichtstrahlen innerhalb der Spezialbildefeffekte. Die folgenden Schritte werden durchgeführt:

- 1) Im ersten Schritt werden die aufgenommenen Bilder rektifiziert. Das heißt, es werden Identifikationspunkte wie Kanten oder Ecken ermittelt und die Bilder so korrigiert, dass diese Identifikationspunkte für die horizontal angebrachten Kameras auf einer horizontalen Linie liegen und für die Punkte von vertikal angeordneten Kameras auf einer vertikalen Linie. Die Positionsverschiebung der Punkte innerhalb der Bilder zum Rand bestimmt die Disparität und damit die Entfernung des Objektes aus diesem Identifikationspunkt zu der Kamera.

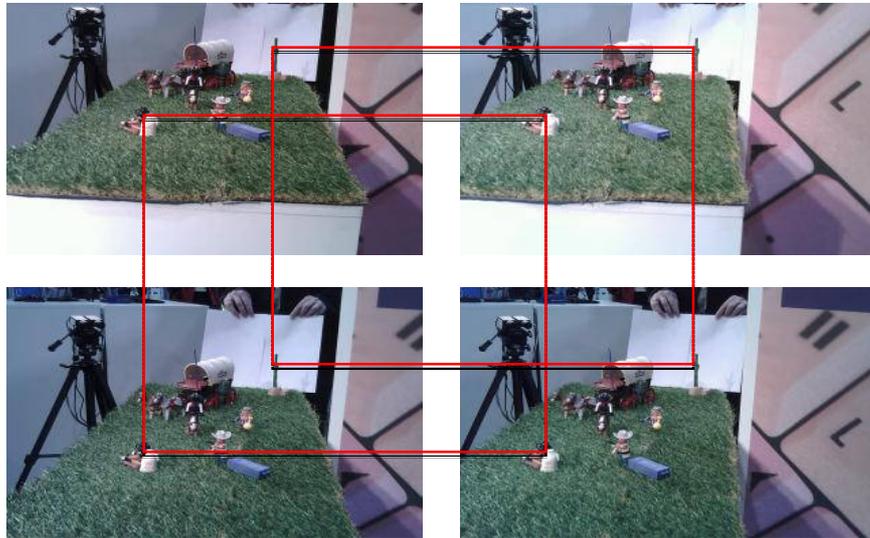


Abb. 6 Auffinden von Korrespondenzpunkten zur Rektifizierung der Bilder

- 2) Der zweite Schritt ist die Berechnung einer Disparitätskarte. Dadurch sind auch unmittelbar Tiefenkarten bestimmt.
- 3) Der dritte Schritt ist die Generierung von Zwischenbildern durch Interpolation (die Disparitätskarten dienen der Interpolation der Zwischenbilder). Hierdurch entsteht ein dichtes Lichtfeld. Je mehr Zwischenbilder berechnet werden, desto dichter ist das Lichtfeld.



Abb. 7 Bilder von realen Kameras (grau umrahmt) und interpolierten Zwischenbildern

Nach dieser Vorverarbeitung können nun die visuellen Spezialeffekte errechnet werden.

## 5 Visuelle Spezialeffekte

Nachdem nun ein dichtes Lichtfeld errechnet wurde, können die Lichtstrahlen in individueller Weise wieder zusammengeführt werden.

### 5.1 *Rendern in z-Richtung (auf die Szene zu)*

Ein erster Effekt ist die Berechnung einer virtuellen Dolly-Fahrt. Um diesen Effekt zu erreichen, werden Lichtstrahlen aus unterschiedlichen Bildern zu einem neuen Bild zusammengesetzt. Die Position der virtuellen Kamera in  $z$ -Richtung definiert, welche Bildteile aus welchen Bildern entnommen werden. Im Gegensatz zu einem rein digitalen Zoom handelt es sich um eine echte Dolly-Fahrt, d.h. auch Verdeckungen und Aufdeckungen ändern sich.

### 5.2 *„Vertigo“: Dolly + Zoom*

Ein zweiter Effekt ist der Vertigo-Effekt.<sup>2</sup> Dabei wird während der Dolly-Fahrt der Zoom verändert, sodass sich ein interessierendes Objekt nicht in der Größe ändert.. Für diesen Effekt müssen die Lichtstrahlen zusätzlich in Abhängigkeit von der Tiefe der einzelnen Szenenteile ausgewählt werden.

### 5.3 *Refokussierung*

Der dritte Effekt ist der klassische Effekt, der gerne zur Demonstration von Lichtfeldkameras eingesetzt wird. Hierbei kann die Tiefenschärfe nach eigenen Wünschen verändert werden – je nachdem, wie viele Lichtstrahlen aus wie vielen Kamerabildern vermischt werden. Hinweis: Bei der Aufnahme einer Szene mit einem Kamera-Array ist das Ziel, eine möglichst große Tiefenschärfe zu verwenden, d.h. die Lichtstrahlen sollten möglichst aus einer Richtung kommen.

---

<sup>2</sup> benannt nach dem Film von Alfred Hitchcock, in dem der Effekt das erste Mal verwendet wurde

#### 5.4 *Free View-point*

Innerhalb der Dimension des physikalischen Kamera-Arrays kann der Blickwinkel verändert werden. Je größer die Dimension, desto größer der Freiheitsgrad. Allerdings heißt das bei gleichbleibender Kamera-Anzahl, dass das Lichtfeld dann auch immer ungenauer erfasst wird und Artefakte entstehen können.

#### 5.5 *3D-Kamera-Berechnung*

Natürlich kann nicht nur eine virtuelle Kamera berechnet werden, sondern auch mehrere. Dies ist insbesondere bei der Berechnung zu stereoskopischen 3D-Filmen von Vorteil. Ähnlich wie im CGI-Bereich können nun die Kameras virtuell positioniert und berechnet werden.

## 6 Lichtfeld Stop-Motion-Film-Test

Der erste Prototyp eines Kamera-Arrays V1 zur Lichtfeldaufzeichnung basiert auf Smartphone-Komponenten. Dieser besitzt daher nur die Fähigkeit der geringen Bildrataufzeichnung. Um trotzdem die Algorithmen und Spezialeffekte auch im Bewegtbild testen zu können, wurde ein Feldversuch auf Basis eines Stop-Motion-Films durchgeführt. Für den Prototyp wurden 16 Kameras in einem 2D-Array angeordnet. Die Dimension des Arrays war ca.  $30 \times 30$  cm.



Abb. 8 Stop-Motion-Setup

Während der gesamten Aufnahme aller Bilder für den Stop-Motion-Film blieb das Kamera-Array unbewegt. Alle Blickwinkel- und Fokusänderungen wurden nachträglich in der Postproduktion errechnet und durchgeführt.

Für die Arbeit mit dem Kamera-Array wurden verschiedene Software-Werkzeuge entwickelt:

- Ein Kamera-Arrayviewer, der die Bilder aller 16 Kameras als Preview-Thumbnail anzeigt sowie ein ausgewähltes Bild in voller Auflösung: Dies ermöglicht die Kontrolle der Kameras und des Blickfeldraumes – und damit eine gute Einstellmöglichkeit.
- Ein Lichtfeld Bildvorrender-Programm: Dieses errechnet nach Auswahl auf einem  $xy$ -Feld eine virtuelle Kameraposition. Ebenso ist über einen Schieberegler die Verschiebung der Fokusebene möglich. Damit ist eine Beurteilung der möglichen Freiheitsgrade gegeben.

Basierend auf diesen Werkzeugen wurde der Stop-Motion-Film aufgenommen und eine virtuelle Dolly-Fahrt mit kombinierter Schärfeverlagerung errechnet.

## 7 Ausblick

Die ersten Tests zeigten, dass das Konzept der Lichtfeldaufnahme mit nachgelagerter Blickwinkel- und Fokuseinstellung funktioniert. Verbesserungen, insbesondere in der Detektion von Identifikationspunkten, sind noch notwendig, um die Qualität der gerenderten Bilder zu verbessern.

Hauptschwerpunkte der weiteren Entwicklung sind der Einsatz von HD-Kameras mit 30 bzw. 60 Bildern pro Sekunde. Besondere Aufmerksamkeit wird hierbei auf die Synchronisation der einzelnen Kameras untereinander und zu externen Fremdkameras zu legen sein. Ein spezieller Entwicklungsschwerpunkt ist auch das intelligente Datenmanagement der riesigen Datenmengen. Dies wird erreicht durch eine Mischung aus Onboard-Speicherung und teilweiser Übertragung der Daten zu einem Zentralrechner über Ethernet. Eine Integration der Algorithmen als Plug-in in kommerzielle Postproduktionssoftware wird angestrebt.

## 8 Zusammenfassung

Das realisierte Konzept zeigt, dass ein Multi-Kamera-Array ein dünn besetztes Lichtfeld aufzeichnen kann. Mithilfe zusätzlicher Vorverarbeitungsschritte kann ein dicht besetztes Lichtfeld errechnet werden. Basierend auf dem

dicht besetzten Lichtfeld können visuelle Spezialeffekte in der Postproduktion eingesetzt werden, die normalerweise fest während der Aufnahme ‚eingebakken‘ sind. Dazu gehören die Wahl eines freien Blickwinkels, eine nachträgliche Refokussierung oder komplexere Effekte wie der „Vertigo“- oder „Matrix“-Effekt. Ein interessanter Aspekt ist auch, mehr als eine virtuelle Kamera zu berechnen, sodass für 3D-Filme sehr einfach verschiedene Parallaxen berechnet werden können.

Je größer die Anzahl der Kameras, desto besser ist die Bildqualität auf den final berechneten Bildern, da mehr Interpolationspunkte für die Berechnung des dichten Lichtfeldes zur Verfügung stehen. Je größer die Dimension des Kamera-Arrays, desto größer sind die Freiheitsgrade in der Positionierung der virtuellen Kameras.

Besondere Aufmerksamkeit muss der Synchronisation der Kameras und dem Datenfluss gewidmet werden. Mit 10 oder evtl. sogar einmal 100 Kameras pro Array entstehen riesige Datenmengen. Daher benötigt das Konzept ein intelligentes und optimiertes Datenflussmanagement. Auf der anderen Seite sind hierdurch neue Freiheitsgrade in der gestalterischen Arbeit möglich.

## Weiterführende Literatur

- M. Faraday (1846): Thoughts on Ray Vibrations. In: *Philosophical Magazine*, Bd. XXVIII, N188, S. 3.
- T. Georgiev; A. Lumsdaine (2010): Focused plenoptic camera and rendering. In: *Journal of Electronic Imaging* 19 (2).
- A. Gershun (1936): *The Light Field*. Moskau (übers. von P. Moon und G. Timoshenko, in: *Journal of Mathematics and Physics* Vol. XVIII [1939], S. 51–151).
- M. Levoy; P. Hanrahan (1996): Light Field Rendering. In: *Proc. ACM SIGGRAPH*. ACM Press, S. 31–42.
- G. Lippman (1908): La Photographie Intégrale. In: *Comptes-Rendus, Académie des Sciences* 146, S. 446–551.