

Meine erste pauschale Anfrage an ChatGPT:

Ich habe die Idee, etwa 50 (Street) Fotografien die ich als Bildband zusammengestellt habe von einer KI analysieren zu lassen. Aber nicht um ihren Inhalt direkt zu beschreiben und auch nicht um daraus neue Bilder zu generieren, sondern um eine gewisse Stimmung zu ermitteln, die diese Bilder und die Art ihrer Präsentation vermitteln. Als Ergebnis soll ein etwa 2 DIN A4 Seiten umfassender Text generiert werden, der als Epilog in diesem Bildband abgedruckt wird. Der Text soll wie gesagt nicht direkten Bezug zu den Bildern haben. Es sollte eine leicht surreale Geschichte werden, die nur ein einziges Mal den Titel des Bildbandes in Großbuchstaben enthält. Denkst du, dass so etwas möglich ist und wie und mit welchen Werkzeugen soll ich vorgehen?

---

Ja — das ist nicht nur möglich, sondern konzeptionell sehr gut geeignet für KI. Dein Ansatz ist sogar ausgesprochen *zeitgemäß* für einen Bildband: Die KI fungiert weniger als „Beschreiberin“ der Bilder, sondern als **Resonanzraum** für Stimmung, Rhythmus und Präsentation.

Ich skizziere dir erst **warum das funktioniert**, dann **wie du konkret vorgehen kannst**, und schließlich **welche Werkzeuge sinnvoll sind**.

---

## 1. Warum das funktioniert

KI-Modelle (insbesondere multimodale Sprachmodelle) sind erstaunlich gut darin,

- **latente Stimmungen** aus Bildserien abzuleiten  
(z. B. Enge vs. Weite, Beschleunigung vs. Stillstand, Anonymität, Melancholie, Reibung)
- **Muster in Präsentation** zu erfassen  
(Sequenzierung, Wiederholungen, Brüche, Leerstellen, Seitenrhythmus)
- daraus **nicht-referenzielle Texte** zu generieren  
→ also Texte, die *nicht* sagen: „Auf dem Bild sieht man ...“, sondern eher:  
*„Es gibt Orte, an denen niemand stehen bleibt, obwohl alle dort gewesen sind.“*

Gerade Street Photography eignet sich hervorragend, weil sie ohnehin mit Andeutung, Fragment und Zufälligkeit arbeitet. Die KI kann das auf eine narrative Ebene übertragen, ohne die Bilder zu verraten oder zu erklären.

---

## 2. Empfohlener Arbeitsprozess (sehr wichtig)

### Schritt 1: Bildband als System denken

Nicht einzelne Bilder analysieren lassen, sondern:

- Reihenfolge festlegen (final!)
- Doppelseiten / Leerraum / Kapiteltrennungen berücksichtigen
- evtl. 5–7 „Cluster“ definieren (ohne sie zu benennen)

Die KI reagiert stärker auf **Sequenzen** als auf Einzelbilder.

---

## Schritt 2: Bildmaterial übergeben

Optimal wäre:

- alle ~50 Bilder
- in der finalen Reihenfolge
- ggf. als PDFs oder nummerierte JPEGs

Du gibst **keine inhaltlichen Erklärungen** zu den Bildern ab.

---

## Schritt 3: Analyse-Prompt (nicht Text-Prompt!)

Zuerst **keinen literarischen Text verlangen**, sondern z. B.:

Analysiere diese Bildserie ausschließlich auf

- Atmosphäre
- emotionalen Grundton
- zeitliches Gefühl
- Verhältnis von Mensch / Raum / Bewegung
- Wirkung der Sequenzierung

Beschreibe keine Motive, keine Orte, keine erkennbaren Inhalte.

Das Ergebnis ist eine **Stimmungs-Matrix**, kein Text für den Band.

---

## Schritt 4: Transformation in eine surreale Geschichte

Erst *danach* gibst du einen zweiten Auftrag:

- Textform: leicht surreal, episodisch oder fließend
- Länge: ca. 2 DIN-A4-Seiten
- Ton: ruhig, beobachtend, leicht entrückt
- **explizite Regeln:**
  - keine direkten Bildbeschreibungen
  - keine Ortsnamen
  - keine Kamera-/Fotografie-Referenzen
  - **Titel des Bildbands genau ein Mal in GROSSBUCHSTABEN**
  - Text soll als *Epilog* funktionieren, nicht als Erklärung

Das ist entscheidend:

👉 Die KI schreibt *aus der extrahierten Stimmung*, nicht aus den Bildern selbst.

---

## Schritt 5: Kuratieren (nicht blind übernehmen)

Sehr wahrscheinlich wirst du:

- 5–15 % kürzen
- einzelne Sätze verschieben
- evtl. das Ende leicht anpassen

Das ist keine Schwäche, sondern Teil der Autorenschaft.

---

## 3. Geeignete Werkzeuge

### ✅ Am besten geeignet

- **ChatGPT mit Bildverarbeitung**  
(multimodal, sehr gut für atmosphärische Ableitungen)
- Alternativ: **Claude (Anthropic)**  
sehr stark in literarischem Ton und subtiler Narration

### ⚠️ Weniger geeignet

- reine Bildanalyse-Tools (z. B. automatische Tagger)
  - Bildgeneratoren (Midjourney, DALL-E etc.) → falscher Fokus
  - klassische Text-KIs *ohne* Bildinput
- 

## 4. Künstlerisch gedacht: Was du hier eigentlich tust

Du lässt die KI nicht „über deine Bilder schreiben“, sondern:

**Du benutzt sie als eine fremde Wahrnehmung,**  
die zwischen Betrachter und Werk geschaltet ist.

Der Epilog wird damit:

- kein Kommentar
- keine Interpretation
- sondern eine **parallele Realität**, die aus derselben Stimmung geboren wurde

Das ist extrem passend für Street Photography.

---

## 5. Wenn du willst, kann ich dir helfen bei ...

- der **Formulierung der Analyse-Prompts**
- einer **Beispielstruktur für den Epilog**
- der Entscheidung: *surreale Kurzgeschichte vs. fragmentarischer Text vs. lyrische Prosa*
- einem **Testlauf mit 5–7 Bildern**, bevor du alles einreichst

Wenn du magst, sag mir:

- ob der Bildband eher **dunkel / leise / roh / ironisch / melancholisch** ist
- ob du eher **Ich-Erzähler, unpersönliche Stimme** oder **kollektives „Wir“** willst

Dann können wir das sehr präzise aufsetzen.

Vielen Dank! Das ist ein ausgezeichneter Vorschlag! Ich werde das in Angriff nehmen. Sag mir bitte, wie ich die Bilder - aber letztendlich den Bildband in seiner Gänze und Zusammenstellung an ChatGPT übertrage.

Sehr gern — dann gehen wir jetzt **ganz praktisch** vor.

Ich erkläre dir **konkret**, wie du den Bildband an ChatGPT übergibst, **in welcher Form, in welcher Reihenfolge**, und **wie du sicherstellst, dass die KI „den Band als Ganzes“ versteht** und nicht nur einzelne Bilder.

---

### Grundsatz (wichtig)

ChatGPT kann **keine echten Bücher „durchblättern“**, aber es kann:

- ✓ eine **Sequenz von Bildern als kohärentes Werk** erfassen
- ✓ Seitenrhythmus, Wiederholung und Brüche mitdenken
- ✓ aus der Gesamtheit eine **übergeordnete Stimmung** abstrahieren

Das erreichst du durch **strukturierte Übergabe**, nicht durch Kommentare.

---

### OPTION A (empfohlen): Upload als sequenzierte Einzelbilder

#### Schritt 1: Bilder vorbereiten

- Exportiere die finalen Bilder als:
  - JPG oder PNG
  - gleiche Seitenproportion (z. B. Hochformat A4 oder quadratisch)
- Benenne sie strikt numerisch:
  - 01.jpg
  - 02.jpg
  - 03.jpg
  - ...
  - 50.jpg

## 👉 Keine Titel, keine Metadaten, keine Beschreibungen

---

### Schritt 2: Upload in ChatGPT

- Ziehe alle Bilder **in einer einzigen Nachricht** in den Chat (oder in 2–3 Nachrichten, falls es technisch limitiert ist)
  - **Nicht zwischendurch kommentieren**
- 

### Schritt 3: Einleitender Systemtext (sehr wichtig)

Schreibe **vor oder nach dem Upload** genau einmal z. B.:

Die folgenden Bilder bilden einen abgeschlossenen Bildband.

Die Reihenfolge ist final und bewusst gewählt.

Bitte betrachte sie als fortlaufende Sequenz, nicht als Einzelwerke.

Es sollen keine einzelnen Bilder beschrieben oder benannt werden.

Das „programmiert“ den Blick der KI.

---

### OPTION B: Upload als PDF (funktioniert, aber weniger fein)

#### Wann sinnvoll?

- Wenn du Doppelseiten, Leerraum oder Kapiteltrennungen zeigen willst
- Wenn Seitenverhältnisse wichtig sind

#### Vorgehen

- Erstelle ein **einfaches PDF**
  - eine Seite = eine Buchseite
  - ggf. leere Seiten explizit drin lassen
- Keine Texte, keine Seitennummern

👉 Nachteil: ChatGPT erkennt zwar Sequenz, aber weniger gut einzelne Übergänge.

---

### OPTION C (ideal, aber optional): Kombination

Wenn du sehr präzise sein willst:

1. **Einzelbilder nummeriert** (für Bildinhalt & Stimmung)
2. **Zusätzlich ein PDF** (für Rhythmus & Layoutgefühl)

Dann schreibst du:

Die Einzelbilder und das PDF zeigen dasselbe Werk.

Nutze die Bilder für Atmosphäre, das PDF für Sequenzgefühl.

---

## Der entscheidende Prompt (bitte so oder sehr ähnlich verwenden)

### Prompt 1: Analyse (noch kein Text!)

Nachdem die Bilder hochgeladen sind:

Analysiere diesen Bildband ausschließlich auf:

- atmosphärischen Grundton
- emotionale Spannung über die Sequenz hinweg
- implizites Zeitgefühl
- Verhältnis von Nähe und Distanz
- Wirkung von Wiederholung, Verdichtung und Leerräumen

Beschreibe keine Motive, keine Orte, keine Personen.

Vermeide jede direkte Referenz auf einzelne Bilder.

Formuliere das Ergebnis als zusammenhängende, abstrahierende Analyse.

👉 Das Ergebnis darf ruhig „essayistisch“ sein, aber **nicht literarisch**.

---

### Prompt 2: Transformation in den Epilog

Erst nach der Analyse:

Verwandle die ermittelte Stimmung und Dynamik des Bildbands in eine leicht surreale, prosaische Erzählung.

Bedingungen:

- Länge ca. 2 DIN-A4-Seiten
- epilogischer Charakter
- keine direkte Bezugnahme auf Bilder, Fotografie, Orte oder reale Ereignisse
- keine erklärende oder kommentierende Stimme
- der Titel **[DEIN TITEL]** soll genau ein einziges Mal in GROSSBUCHSTABEN erscheinen
- ruhiger, schwebender Ton mit unterschwelliger Spannung

Der Text soll wirken, als existiere er parallel zu den Bildern, nicht als ihr Kommentar.

---

### Woran du erkennst, dass es funktioniert hat

Der Text ist gelungen, wenn:

- Leser:innen **ohne Bilder** etwas spüren
- Leser:innen **mit Bildern** das Gefühl haben:  
„Das ist nicht über die Bilder geschrieben — aber es passt unheimlich gut.“
- nichts im Text „entlarvt“, was konkret abgebildet ist

---

### Letzter, wichtiger Hinweis (künstlerisch)

Du bleibst Autor.

Die KI ist **nicht der Verfasser**, sondern:

eine fremde Stimme, die du eingeladen hast,  
einmal durch dein Werk hindurchzugehen.

Wenn du willst, können wir:

- vorab einen **Test mit 5 Bildern** machen
- deinen Prompt **feinschleifen**
- oder nach dem ersten Text eine **zweite, radikalere Version** erzeugen

Sag einfach Bescheid — und viel Erfolg, das ist ein starkes Projekt.