PTO/AIA/15 (10-17)
Approved for use through 05/31/2024. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	PP6807264
First Named Inventor	Sanghyun PARK
Title	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSI
Priority Mail Express® Label No.	

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. [✔] **Fee Transmittal Form** (PTO/SB/17 or equivalent)
2. [✔] **Applicant asserts small entity status.** See 37 CFR 1.27
3. [ ] **Applicant certifies micro entity status.** See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.
4. [✔] **Specification** [*Total Pages* 32] Both the claims and abstract must start on a new page. *(See MPEP § 608.01(a) for information on the preferred arrangement)*
5. [✔] **Drawing(s)** (35 U.S.C. 113) [*Total Sheets* 10]
6. **Inventor's Oath or Declaration** [*Total Pages* 4] *(including substitute statements under 37 CFR 1.64 and assignments serving as an oath or declaration under 37 CFR 1.63(e))*
   a. [✔] Newly executed (original or copy)
   b. [ ] A copy from a prior application (37 CFR 1.63(d))
7. [✔] **Application Data Sheet** ** See note below.* See 37 CFR 1.76 (PTO/AIA/14 or equivalent)
8. **CD-ROM or CD-R** in duplicate, large table, or Computer Program (*Appendix*)
   [ ] Landscape Table on CD
9. **Nucleotide and/or Amino Acid Sequence Submission** (*if applicable, items a. – c. are required*)
   a. [ ] Computer Readable Form (CRF)
   b. [ ] Specification Sequence Listing on:
      i. [ ] CD-ROM or CD-R (2 copies); or
      ii. [ ] Paper
   c. [ ] Statements verifying identity of above copies

ADDRESS TO:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

ACCOMPANYING APPLICATION PAPERS

10. [ ] **Assignment Papers** (cover sheet & document(s))
    Name of Assignee _____
11. [ ] **37 CFR 3.73(c) Statement** (*when there is an assignee*) [✔] **Power of Attorney**
12. [ ] **English Translation Document** (*if applicable*)
13. [✔] **Information Disclosure Statement** (PTO/SB/08 or PTO-1449)
    [✔] Copies of citations attached
14. [ ] **Preliminary Amendment**
15. [ ] **Return Receipt Postcard** (*MPEP § 503*) (*Should be specifically itemized*)
16. [ ] **Certified Copy of Priority Document(s)** (*if foreign priority is claimed*)
17. [ ] **Nonpublication Request** Under 35 U.S.C. 122(b)(2)(B)(i). Applicant must attach form PTO/SB/35 or equivalent.
18. [✔] **Other:** Request to retrieve electronic priority application; Authorization for internet communication

Note: (1) Benefit claims under 37 CFR 1.78 and foreign priority claims under 1.55 **must** be included in an Application Data Sheet (ADS).
(2) For applications filed under 35 U.S.C. 111, the application must contain an ADS specifying the applicant if the applicant is an assignee, person to whom the inventor is under an obligation to assign, or person who otherwise shows sufficient proprietary interest in the matter. See 37 CFR 1.46(b).

19. CORRESPONDENCE ADDRESS

[✔] The address associated with Customer Number: 183715 **OR** [ ] Correspondence address below

Name					
Address					
City		State		Zip Code	
Country		Telephone		Email	

Signature	/Byungwoong Park/	Date	2025-03-13
Name (Print/Type)	Byungwoong Park	Registration No. (Attorney/Agent)	80,134

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. The United States Patent and Trademark Office (USPTO) collects the information in this record under authority of 35 U.S.C. 2. The USPTO's system of records is used to manage all applicant and owner information including name, citizenship, residence, post office address, and other information with respect to inventors and their legal representatives pertaining to the applicant's/owner's activities in connection with the invention for which a patent is sought or has been granted. The applicable Privacy Act System of Records Notice for the information collected in this form is COMMERCE/PAT-TM-7 Patent Application Files, available in the Federal Register at 78 FR 19243 (March 29, 2013). https://www.govinfo.gov/content/pkg/FR-2013-03-29/pdf/2013-07341.pdf

Routine uses of the information in this record may include disclosure to: 1) law enforcement, in the event that the system of records indicates a violation or potential violation of law; 2) a Federal, state, local, or international agency, in response to its request; 3) a contractor of the USPTO having need for the information in order to perform a contract; 4) the Department of Justice for determination of whether the Freedom of Information Act (FOIA) requires disclosure of the record; 5) a Member of Congress submitting a request involving an individual to whom the record pertains, when the individual has requested the Member's assistance with respect to the subject matter of the record; 6) a court, magistrate, or administrative tribunal, in the course of presenting evidence, including disclosures to opposing counsel in the course of settlement negotiations; 7) the Administrator, General Services Administration (GSA), or their designee, during an inspection of records conducted by GSA under authority of 44 U.S.C. 2904 and 2906, in accordance with the GSA regulations and any other relevant (i.e., GSA or Commerce) directive, where such disclosure shall not be used to make determinations about individuals; 8) another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)); 9) the Office of Personnel Management (OPM) for personnel research purposes; and 9) the Office of Management and Budget (OMB) for legislative coordination and clearance.

If you do not furnish the information requested on this form, the USPTO may not be able to process and/or examine your submission, which may result in termination of proceedings, abandonment of the application, and/or expiration of the patent.

Additional Uses

Additional USPTO uses of the information in this record may include disclosure to: 1) the International Bureau of the World Intellectual Property Organization, if the record is related to an international application filed under the Patent Cooperation Treaty; 2) the public i) after publication of the application pursuant to 35 U.S.C. 122(b), ii) after issuance of a patent pursuant to 35 U.S.C. 151, iii) if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections, or an issued patent, or iv) without publication of the application or patent under the specific circumstances provided for by 37 CFR 1.14(a)(1)(v)-(vii); and/or 3) the National Archives and Records Administration, for inspection of records.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Sanghyun		PARK	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence[i]	KR

Mailing Address of Inventor:

Address 1	101-1502, 30 Semuseo 8-gil, Seodaemun-gu		
Address 2			
City	Seoul	State/Province	
Postal Code	03626	Country[i]	KR

Inventor 2 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Kijung		LEE	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence[i]	KR

Mailing Address of Inventor:

Address 1	G-201, 24-4 Yeonhui-ro 10-gil, Seodaemun-gu		
Address 2			
City	Seoul	State/Province	
Postal Code	03725	Country[i]	KR

Inventor 3 Remove

Legal Name

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Prefix	Given Name	Middle Name	Family Name	Suffix
	Youngwan		JO	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence[i]	KR

Mailing Address of Inventor:

Address 1	#402, 5 Seongsan-ro 17-gil, Seodaemun-gu		
Address 2			
City	Seoul	State/Province	
Postal Code	03726	Country[i]	KR

Inventor 4 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Sunghyun		AHN	

Residence Information (Select One) ○ US Residency ◉ Non US Residency ○ Active US Military Service

City	Seoul	Country of Residence[i]	KR

Mailing Address of Inventor:

Address 1	104-1901, 15 Wonjeok-ro 269beon-gil, Bupyeong-gu		
Address 2			
City	Incheon	State/Province	
Postal Code	21368	Country[i]	KR

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ **An Address is being provided for the correspondence Information of this application.**

Customer Number	183715		
Email Address	bwp@zionip.com	Add Email	Remove Email
Email Address	patent_01@zionip.com	Add Email	Remove Email
Email Address	pbwoong@hotmail.com		Remove Email

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Application Information:

Title of the Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR		
Attorney Docket Number	PP6807264	**Small Entity Status Claimed**	☒
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	10	**Suggested Figure for Publication (if any)**	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	◉ Customer Number	○ US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	183715		

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status			Remove
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button.

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) the[i] information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

			Remove
Application Number	Country[j]	Filing Date (YYYY-MM-DD)	Access Code[i] (if applicable)
10-2024-0055081	KR	2024-04-25	7E74

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Applicant 1

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

(●) Assignee	( ) Legal Representative under 35 U.S.C. 117	( ) Joint Inventor
( ) Person to whom the inventor is obligated to assign.	( ) Person who shows sufficient proprietary interest	

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. [X]

Organization Name	UIF (University Industry Foundation), Yonsei University

Mailing Address Information For Applicant:

Address 1	50 Yonsei-ro, Seodaemun-gu		
Address 2			
City	Seoul	**State/Province**	
Country	KR	Postal Code	03722
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button.

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	**Given Name**	Middle Name	**Family Name**	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		**State/Province**	
Country[i]		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Byungwoong Park/			Date (YYYY-MM-DD)	2025-03-13
First Name	Byungwoong	Last Name	Park	Registration Number	80,134

Additional Signature may be generated within this form by selecting the Add button.

PTO/AIA/14 (02-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	PP6807264
	Application Number	
Title of Invention	METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR	

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1 The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3 A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent CooperationTreaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Doc Code: ECOMM.AUTH/ECOMM.WTDW
Doc Description: Internet Communications Authorization/Internet Communications Authorization Withdrawal

PTO/SB/439 (11-15)

AUTHORIZATION FOR INTERNET COMMUNICATIONS IN A PATENT APPLICATION OR REQUEST TO WITHDRAW AUTHORIZATION FOR INTERNET COMMUNICATIONS		
	Application No.	
	Filing Date	2025-03-13
	First Named Inventor	Sanghyun PARK
	Art Unit	
	Examiner Name	
	Practitioner Docket No.	PP6807264

To: Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I. To authorize permission for Internet Communications.

☒ Recognizing that Internet communications are not secure, I hereby authorize the USPTO to communicate with the undersigned and practitioners in accordance with 37 CFR 1.33 and 37 CFR 1.34 concerning any subject matter of this application via video conferencing, instant messaging, or electronic mail. I understand that a copy of these communications will be made of record in the application file. (MPEP 502.03)

II. To withdraw authorization for Internet Communications.

☐ The authorization given on _____, to the USPTO to communicate with the undersigned and any practitioner in accordance with 37 CFR 1.33 and 37 CFR 1.34 concerning any subject matter of this application via Internet communications such as video conferencing, instant messaging, or electronic mail is hereby withdrawn. I understand that the withdrawal is effective when approved rather than when received.

I am the

☐ applicant.

☐ attorney or agent of record. Registration number _____.

☒ attorney or agent acting under 37 CFR 1.34. Registration number 80134.

/Byungwoong Park/ — Signature

2025-03-13 — Date

Byungwoong Park — Typed or printed name

301-246-0134 — Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Juristic entities must be represented by a patent practitioner (see 37 CFR 1.31, which is applicable to any paper filed on or after September 16, 2012, that is presented on behalf of a juristic entity, regardless of application filing date). Submit multiple forms if more than one signature is required, see below*.

☐ * Total of _____ forms are submitted.

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Doc Code: PD.REQ.RETR
Document Description: Request for USPTO to retrieve priority docs

PTO/SB/38 (08-22)
Approved for use through 05/31/2024. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Request to Retrieve Electronic Priority Application(s)

Send completed form to: Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

COMPLETE IF KNOWN	
Application Number	
Filing Date	2025-03-13
First Named Inventor	Sanghyun PARK
Art Unit	
Examiner Name	
Attorney Docket Number	PP6807264

Pursuant to 37 CFR 1.55(i), the undersigned hereby requests that the USPTO retrieve an electronic copy of each of the following foreign applications for which priority has been claimed under 35 U.S.C. 119(a)-(d) from a foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement. This Request must be submitted:

- within the later of sixteen months from the filing date of the prior foreign application or four months from the actual filing date of an application under 35 U.S.C. 111(a),
- within four months from the later of the date of commencement (37 CFR 1.491(a)) or the date of the initial submission under 35 U.S.C. 371 of an application entering the national stage under 35 U.S.C. 371, or
- with a petition under 37 CFR 1.55(e) or (f).

☐ **OPTION A**

Please retrieve the priority application identified in Column C, a certified copy of which is contained in the EP or JP application identified in Columns A and B:

	A	B			C	
	Code for Participating Office (EP or JP only)	Application containing the non-participating priority application			Non-participating priority application to be retrieved	
		App. No.	Filing Date	Access Code	Country Code	App. No.
1						

☑ **OPTION B**

This Request may be used for the infrequent circumstance when a claim for priority to an application filed in a participating foreign intellectual property office was made prior to that foreign intellectual property office becoming a participating foreign intellectual property office.

Please retrieve the priority application identified in Columns A and B:

	A	B		
	Code for Participating Office (e.g., EP) or WIPO DAS Depositing Office (e.g., AU, BR, CN, DK, EA, EE, ES, FI, GB, IB, IN, JP, KR, MA, NL, NZ, SE)	Application to be retrieved		
		App. No.	Filing Date	Access Code (for WIPO DAS Depositing Office)
1	KR	10-2024-0055081	2024-04-25	7E74
2				

The USPTO will not attempt to retrieve the identified priority application(s) unless an identical claim for foreign priority to the application identified above is made pursuant to 37 CFR 1.55(d) or a petition is granted under 37 CFR 1.55(e) or (f). Applicants are advised to consult Private PAIR (accessed through www.uspto.gov) to assure that the retrieval has been successful. The applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period set forth in 37 CFR 1.55(g)(1).

I hereby declare that I have the authority to grant access to the above-identified foreign application(s).

/Byungwoong Park/	2025-03-13
Signature	Date
Byungwoong Park	301-246-0134
Printed or Typed Name	Telephone Number
Attorney of Record	80,134
Title	Registration Number, if applicable

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

PTO/SB/08a (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(**Not for submission under 37 CFR 1.99)**

Application Number	
Filing Date	2025-03-13
First Named Inventor	Sanghyun PARK
Art Unit	
Examiner Name	
Attorney Docket Number	PP6807264

U.S.PATENTS

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Patent citation information please click the Add button.

U.S.PATENT APPLICATION PUBLICATIONS

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Published Application citation information please click the Add button.

FOREIGN PATENT DOCUMENTS

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	10-2022-0088653	KR	A	2022-06-28	AHA Co., Ltd.		☒
	2	2018-077479	JP	A	2018-05-17	QUALCOMM INCORPORATED		☒
	3	10-2323671	KR	B1	2021-07-08	INDUSTRY ACADEMY COOPERATION FOUNDATION et al.		☒

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	
	Filing Date	2025-03-13
	First Named Inventor	Sanghyun PARK
	Art Unit	
	Examiner Name	
	Attorney Docket Number	PP6807264

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1	Anonymous ECCV et al. "MDVAD: Multimodal Diffusion for Video Anomaly Detection", 2024.	☐
	2	Jing Ren et al. "Deep Video Anomaly Detection: Opportunities and Challenges", 2021-10-11.	☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	
	Filing Date	2025-03-13
	First Named Inventor	Sanghyun PARK
	Art Unit	
	Examiner Name	
	Attorney Docket Number	PP6807264

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Byungwoong Park/	Date (YYYY-MM-DD)	2025-03-13
Name/Print	Byungwoong Park	Registration Number	80,134

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U. S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR

CROSS REFERENCE TO RELATED APPLICATION

The present application claims priority to Korean Patent Application No. 10-2024-0055081, filed April 25, 2024, the entire contents of which are incorporated herein for all purposes by this reference.

STATEMENT REGARDING GOVERNMENT SPONSORED RESEARCH OR DEVELOPMENT

The present disclosure was developed in the task of a project (Project identification number: 1711198526, Project number: 00229822, Ministry name: Ministry of Science and ICT, Project management organization name: National Research Foundation of Korea, Research project name: Innovative New Drug Discovery Using Artificial Intelligence, Research Project Name: Development of an AI-based Multi-drug Indication Optimization Platform and Innovative New Drug Discovery for Overcoming Intractable Diseases, Project implementation organization name: Yonsei University, Research period: 2024.01.01~2024.12.31.)

Meanwhile, in all the aspects of the inventive concept, there is no property interest in the government of the Republic of Korea.

BACKGROUND

Field of the Invention

The present disclosure relates to a method of detecting a video anomaly on the basis of multimodal diffusion and a device therefor and, more particularly, to a method of detecting a video anomaly by using a plurality of features and a device therefor.

Description of the Related Art

Recently, with the development of technologies such as artificial intelligence (AI), various technologies are being developed to recognize abnormal behaviors related to the occurrence of safety accidents, etc. through images collected from surveillance cameras such as CCTV. For example, AI models are being trained and developed to distinguish between images captured in normal conditions and images captured when abnormal behaviors occur. However, since the occurrence frequency of abnormal behaviors is low, it is difficult to secure sufficient image data for training such AI models. In addition, most current models may only utilize fragmentary information such as frame images, resulting in low accuracy.

SUMMARY

An objective of the present disclosure for solving the problem described above is to provide: a method of detecting a video anomaly on the basis of multimodal diffusion; a computer program stored in a computer-readable medium; the computer-readable medium stored with the computer program; and a device (a system) therefor.

According to an exemplary embodiment of the present disclosure, there is provided a method of detecting a video anomaly on the basis of multimodal diffusion and being performed by at least one processor, the method including: obtaining video data including a plurality of frames; detecting an object included in each of the plurality of frames; extracting a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the

detected object; generating a noise vector by injecting noise into the visual feature vector; generating a restoration vector with the noise removed by inputting the noise vector into a diffusion model and using the text feature vector and the motion feature vector as conditions; and performing anomaly detection on the video data by comparing the visual feature vector and the restoration vector.

According to the exemplary embodiment of the present disclosure, the extracting of the multimodal feature vector may include extracting the visual feature vector for the object by providing information related to the detected object to a trained model based on Inflated 3D ConvNet (I3D).

According to the exemplary embodiment of the present disclosure, the extracting of the multimodal feature vector may include generating a caption for describing the object by providing information related to the detected object to a model based on Bidirectional Encoder Representations from Transformers (BERT); and extracting the text feature vector corresponding to the description of the object by providing the generated caption to a trained model based on Simple Contrastive Learning of Sentence Embeddings (SimCSE).

According to the exemplary embodiment of the present disclosure, the extracting of the multimodal feature vector may include extracting skeletal information corresponding to the object by providing information related to the detected object to a trained model based on High-Resolution Network (HRNet); and extracting the motion feature vector representing motion of the object by using the extracted skeletal information.

According to the exemplary embodiment of the present disclosure, the extracting of the motion feature vector representing the motion of the object by using the extracted skeletal information may include extracting the motion feature vector by providing the extracted skeletal information to a trained model based on PoseConv3D.

According to the exemplary embodiment of the present disclosure, the generating of the

noise vector by injecting the noise into the visual feature vector may include generating the noise vector by injecting an amount of Gaussian noise determined according to a range of a time step into the visual feature vector.

According to the exemplary embodiment of the present disclosure, the diffusion model may include a first diffusion model and a second diffusion model, and the generating of the restoration vector with the noise removed may include a first restoration step of inputting the noise vector into the first diffusion model and removing at least some of the noise included in the noise vector by using the text feature vector as a condition; and a second restoration step of inputting a noise vector into the second diffusion model and removing at least some of the noise included in the noise vector by using the motion feature vector as a condition.

According to the exemplary embodiment of the present disclosure, the generating of the restoration vector with the noise removed may further include generating the restoration vector with the noise removed by iteratively performing the first restoration step and the second restoration step.

According to the exemplary embodiment of the present disclosure, the performing of the anomaly detection on the video data may include calculating an anomaly score based on a distance between the visual feature vector and the restoration vector; and performing the anomaly detection on the video data on the basis of whether the calculated anomaly score is greater than or equal to a threshold value.

According to the exemplary embodiment of the present disclosure, the calculating of the anomaly score may include calculating the anomaly score according to the distance by using a mean square error (MSE) between the visual feature vector and the restoration vector.

According to the exemplary embodiment of the present disclosure, the diffusion model may include an encoder including a plurality of denoising attention blocks (DABs); a bottleneck; and a decoder.

According to the exemplary embodiment of the present disclosure, each denoising attention block may include a residual block including a plurality of linear layers connected by skip connection; and a transformer block including a self-attention layer, a cross-attention layer, and a feed-forward network (FFN).

There is provided a computer program stored in a computer-readable recording medium to execute a method, on a computer, described according to the exemplary embodiment of the present disclosure.

According to the exemplary embodiment of the present disclosure, there is provided a computing device including: a communication module; a memory; and at least on processor connected to the memory and configured to execute at least one computer-readable program included in the memory, wherein the at least one program may include commands that obtain video data including a plurality of frames, detect an object included in each of the plurality of frames, extract a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object, generate a noise vector by injecting noise into the visual feature vector, generate a restoration vector with the noise removed by inputting the noise vector into a diffusion model and using the text feature vector and the motion feature vector as conditions, and perform anomaly detection on the video data by comparing the visual feature vector and the restoration vector.

In various exemplary embodiments of the present disclosure, a computing device may enhance the performance of video anomaly detection by complementarily using a multimodal feature vector.

In the various exemplary embodiments of the present disclosure, by referring to a text feature vector and/or a motion feature vector as conditions when a transformer block and a residual block are calculated, a computing device may effectively perform noise removal and vector restoration by referring to both text describing an object and/or motion of the object together with

visual features of the object.

In the various exemplary embodiments of the present disclosure, both a first diffusion model and a second diffusion model having respective conditions different from each other are used instead of using a single diffusion model, so that restoration performance may be improved, and thus video anomaly detection may be performed with higher accuracy.

BRIEF DESCRIPTION OF THE DRAWINGS

The exemplary embodiments of the present disclosure will be described below with reference to the accompanying drawings, wherein similar reference numerals represent similar elements, but are not limited thereto.

FIG. 1 is a functional block diagram illustrating an internal configuration of a computing device according to exemplary embodiments of the present disclosure.

FIG. 2 is a block diagram illustrating a process of extracting a multimodal feature vector according to the exemplary embodiments of the present disclosure.

FIG. 3 is an exemplary view illustrating a structure of a diffusion model according to the exemplary embodiments of the present disclosure.

FIG. 4 is an exemplary view illustrating a structure of a denoising attention block according to the exemplary embodiments of the present disclosure.

FIG. 5 is a view illustrating an example in which a restoration process is performed by a first diffusion model and a second diffusion model according to an exemplary embodiment of the present disclosure.

FIG. 6 is a view illustrating an example in which a restoration process is performed by a first diffusion model and a second diffusion model according to a second exemplary embodiment of the present disclosure.

FIG. 7 is a view illustrating an example in which a restoration process is performed by a first diffusion model and a second diffusion model according to a third exemplary embodiment of the present disclosure.

FIG. 8 is a view illustrating an example in which a restoration process is performed by a first diffusion model and a second diffusion model according to a fourth exemplary embodiment of the present disclosure.

FIG. 9 is a flowchart illustrating an example of a method of detecting a video anomaly on the basis of multimodal diffusion according to the exemplary embodiments of the present disclosure.

FIG. 10 is a block diagram illustrating a hardware configuration of the computing device according to the exemplary embodiments of the present disclosure.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

Hereinafter, specific details for implementing an embodiment of the present disclosure will be described in detail with reference to the attached drawings. However, in the following description, when there is concern of unnecessarily obscuring the gist of the embodiment of the present disclosure, detailed descriptions of well-known functions or components will be omitted.

In the attached drawings, identical or corresponding components are given the same reference numerals. In addition, in the description of the exemplary embodiments below, redundant descriptions of identical or corresponding components may be omitted. However, even though a description of a component is omitted, this omission is not intended to imply that such a component is not included in any exemplary embodiments.

Advantages and features of the disclosed exemplary embodiments and the method of achieving the same will become apparent with reference to the exemplary embodiments described

below in conjunction with the accompanying drawings. However, the present disclosure is not limited to the exemplary embodiments disclosed below, but may be implemented in various different forms. The present exemplary embodiments are provided only to make the present disclosure complete and to fully inform those skilled in the art of the scope of the present disclosure.

The terms used in the present specification will be briefly described, and then the exemplary embodiments of the present disclosure will be described in detail. The terms used in the present specification are the selected general terms that are currently used as widely used as possible while considering functions in the embodiments of the present disclosure, but this may vary according to the intention of those skilled in the art, the judicial precedent, the emergence of new technologies, etc. In addition, in certain cases, there are terms arbitrarily selected by the applicants, and in this case, the meaning of the terms will be described in detail in the description of the corresponding embodiments of the present disclosure. Therefore, the terms used in the present disclosure should be defined based on the meaning of the terms and the overall content of the present disclosure rather than based on simple names of the terms.

In the present specification, singular expressions include plural expressions unless the context clearly specifies that they are singular. In addition, the plural expressions include the singular expressions unless the context clearly specifies that they are plural. Throughout the description of the present specification, when a part is said to "include" or "comprise" a certain component, it means that it may further include or comprise other components, without excluding other components unless the context specifically states otherwise.

In the present disclosure, the terms "comprise," "comprising," and the like may indicate the presence of features, steps, operations, elements, and/or components, but such terms do not exclude the addition of one or more other functions, steps, operations, elements, components, and/or combinations thereof.

In the present disclosure, in a case when a particular component is referred to as being "coupled," "combined," "connected," or "reacting" with any other component, the particular component may be directly coupled, combined, and/or connected to, or reacting with, another component, but is not limited thereto. For example, there may be one or more intermediate components between the particular component and another component. In addition, in the present disclosure, "and/or" may include each of one or more of listed items, or a combination of at least a portion of the one or more of the listed items.

In the present disclosure, terms such as "first," "second," etc. are used to distinguish a particular component from another component, and the components described by such terms are not limited thereto. For example, a "first" component may be an element of the same or similar form as a "second" component.

In the present disclosure, "video anomaly detection" may refer to detecting abnormal behavior and/or abnormal situations such as fights, robberies, arson, explosions, etc., by using images collected from surveillance cameras such as CCTV.

In the present disclosure, "anomaly and/or abnormal behavior" refers to an abnormal behavior predefined by a user, and may include, for example, human action such as fighting, riding a bicycle on a sidewalk, disaster situations such as fire and explosion, and so on.

In the present disclosure, "multimodal" may refer to processing various types of data such as visual data and text data together.

In the present disclosure, a "diffusion model" may refer to a generative model that generates data through a process of gradually adding noise to the data or gradually restoring the data from the noise. For example, the diffusion model may include: a first diffusion model for using a text feature vector as a condition; and a second diffusion model for using a motion feature vector as a condition. Here, the first diffusion model and second diffusion model are trained separately during training, but may be used together during inference.

In the present disclosure, a "visual feature vector" may refer to a vector representing appearance information such as color and shape of an object, a "text feature vector" may refer to a vector representing text describing the object, and a "motion feature vector" may refer to a vector representing a motion of the object. In addition, in the present disclosure, a "noise vector" refers to a vector in which at least some noise is injected into the visual feature vector, and may include both a vector generated by a diffusion process and a vector that has not sufficiently passed through a diffusion model and thus still includes the remaining noise. In addition, in the present disclosure, a "restoration vector" may refer to a vector in a form in which all the noise injected into the visual feature vector is removed.

FIG. 1 is a functional block diagram illustrating an internal configuration of a computing device 100 according to the exemplary embodiments of the present disclosure. According to the exemplary embodiments, the computing device 100, as an arbitrary device for performing video anomaly detection, may include an object detection processor 110, a multimodal feature extraction processor 120, a noise injection processor 130, a vector restoration processor 140, an anomaly detection processor 150, and the like. For example, in a case of obtaining video data including a plurality of frames from a surveillance camera such as CCTV, the computing device 100 may detect whether an abnormal behavior occurs from the corresponding video data.

According to the exemplary embodiments, the computing device 100 may first detect an object included in each of a plurality of frames constituting the corresponding video data in order to detect whether the object included in the video data performs an abnormal behavior. For example, the object detection processor 110 may detect the object included in each of the plurality of frames through any object tracking algorithm (e.g., an object detector, a multi object tracker, etc.) and/or a machine learning model. In this case, an object tracklet as expressed in Equation 1 below may be extracted from the consecutive frames.

[Equation 1]

$$\{O_n | O_n \in \mathbb{R}^{l\times 3\times H\times W}\}_{n=1}^{N}$$

Here, O_n may indicate an object tracklet, N may indicate the number of objects, and L, H, and W may respectively indicate a length, height, and width of the object tracklet. Here, the object tracklet may include an array representing movement over time of one identical object detected on the plurality of frames. That is, the object detection processor 110 may associate the same object extracted from each frame and detect the movement over time of the corresponding object.

According to the exemplary embodiments, the object detection processor 110 may convert the extracted frame-level object tracklet into a segment-level object tracklet. Here, a segment may consist of 16 consecutive frames, but is not limited thereto. In a case of converting the frame-level object tracklet into the segment-level object tracklet, the object tracklet may have a form of $\mathbb{R}^{S\times 16\times 3\times H\times W}$ $S = l/16$. In this way, the segment-level object tracklet converted is information related to the detected object, and may be used as information for multimodal feature extraction.

According to the exemplary embodiments, the multimodal feature extraction processor 120 may extract a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object. For example, the multimodal feature extraction processor 120 may extract a visual feature vector for the object by providing information related to the detected object to a trained model based on Inflated 3D ConvNet (I3D). Here, the I3D-based model may refer to a model for extracting visual information such as color and shape of the object.

Additionally, the multimodal feature extraction processor 120 may generate a caption for providing a description of the object by providing information related to the detected object to a model (e.g., a SwinBERT model) based on Bidirectional Encoder Representations from Transformers (BERT). Here, the caption is a text for describing the detected object. For example, in a case where an object tracklet of "A man riding a bicycle" is provided as input, a caption such as "A man is riding a bicycle with a bicycle on a street" may be extracted. In this case, the multimodal feature extraction processor 120 may extract a text feature vector corresponding to the

description of the object by providing the generated caption to a trained model based on Simple Contrastive Learning of Sentence Embeddings (SimCSE).

Additionally, the multimodal feature extraction processor 120 may extract skeletal information corresponding to the object by providing information related to the detected object to a trained model based on High Resolution Network (HRNet). Here, the skeletal information may be skeletal information generated by extracting key feature points of the object (e.g., joints of a human body, etc.) and connecting the extracted feature points. In this case, the multimodal feature extraction processor 120 may extract a motion feature vector representing motion of the object by providing the extracted skeletal information to a trained model based on PoseConv3D.

According to the exemplary embodiments, the computing device 100 may inject noise onto the visual feature vector in order to use the noise as input to a diffusion model. For example, the noise injection processor 130 may generate a noise vector by injecting an amount of Gaussian noise determined according to a range of a time step into the visual feature vector. The noise injection processor 130 may inject the noise into the visual feature vector on the basis of the following Equation 2 when the time step have a range of $t \in [1, T]$.

[Equation 2]

$$f_{vis}^{t} = \sqrt{\alpha_t} f_{vis}^{0} + \sqrt{1-\alpha_t}\epsilon, \epsilon \sim N(0, I), \alpha_t = \prod_{0}^{t}(1-\beta_t)$$

Here, f_{vis}^{0} indicates a visual feature vector, and f_{vis}^{t} may be a noise vector, i.e., a visual feature vector injected with noise for as long as a time step of t. In addition, β_t may be a schedule used to determine an amount of noise to be injected. That is, as β_t increases, α_t decreases further, so more noise may be injected.

According to the exemplary embodiments, the vector restoration processor 140 may restore an original vector by inputting the noise vector generated by injecting the noise onto the visual feature vector into the diffusion model. For example, the vector restoration processor 140

may input the noise vector into the diffusion model and generate a restoration vector with the noise removed by using a text feature vector and a motion feature vector as conditions. Here, the conditions may refer to information referenced when the diffusion model operates, and the diffusion model may generate data by referencing the information input as the conditions.

According to the exemplary embodiments, the vector restoration processor 140 may generate the restoration vector having the noise removed by iteratively performing restoration steps including: a first restoration step of inputting a noise vector into a first diffusion model and removing at least some of the noise included in the noise vector by using a text feature vector as a condition; and a second restoration step of inputting the noise vector into a second diffusion model and removing at least some of the noise included in the noise vector by using a motion feature vector as a condition.

According to the exemplary embodiments, the anomaly detection processor 150 may perform the anomaly detection on the video data by comparing the visual feature vector and the restoration vector. For example, the anomaly detection processor 150 may calculate an anomaly score based on a distance between the visual feature vector and the restoration vector, and perform the anomaly detection on the video data on the basis of whether the calculated anomaly score is greater than or equal to a threshold value. Here, the anomaly score according to the distance between the visual feature vector and the restoration vector may be calculated by using a mean squared error (MSE) as in the following Equation 3.

[Equation 3]

$$Loss = \left\| f_{vis}^{0} - \hat{f}_{vis}^{0} \right\|_{2}^{2}$$

Here, Loss may indicate a mean square error loss. In addition, f_{vis}^{0} may indicate an initial visual feature vector, and $\hat{f}_{vis}^{0}$ may indicate the restoration vector restored after the noise is removed by the diffusion model.

In FIG. 1, each functional component included in the computing device 100 is separately described, but this is only to help understand the present disclosure, and two or more functions may also be performed in one computing device. With such components, the computing device 100 may improve the performance of video anomaly detection by complementarily using a multimodal feature vector.

FIG. 2 is a block diagram illustrating a process of extracting a multimodal feature vector according to the exemplary embodiments of the present disclosure. As described above, the computing device 100 in FIG. 1 may detect an object in video data and extract a multimodal feature vector including a visual feature vector f_{vis} 212, a motion feature vector f_{mot} 232, and a text feature vector f_{text} 252 for the detected object.

In a case where the object is detected in the video data, the multimodal feature vector may be extracted on the basis of information 202 associated with the detected object. Here, the information 202 associated with the object may represent an object tracklet. According to the exemplary embodiments, in a case where the information 202 associated with the object is input into a visual extractor 210, the visual feature vector f_{vis} 212 may be extracted. Here, the visual extractor 210 is a model configured to recognize and/or classify external information such as color and shape of the object, and may include an I3D-based model, a Convolutional 3D Network (C3D)-based model, etc. For example, the visual feature vector f_{vis} 212 may be extracted as in the following Equation 4.

[Equation 4]

$$f_{vis} = \Phi_{vis}(O_n)$$

According to the exemplary embodiments, skeleton information 222 corresponding to the object may be extracted in a case where the information 202 associated with the object is input into a skeleton extractor 220. Here, the skeleton extractor 220 may include a HRNet-based model for extracting the skeletal information of the object. Additionally, the motion feature vector f_{mot} 232

may be extracted in a case where the skeletal information 222 is input into a motion extractor 230. Here, the motion extractor 230 may include a PoseConv3D-based model and the like for estimating the pose and/or motion of the object on the basis of the skeletal information 222. For example, the motion feature vector f_{mot} 232 may be extracted as in the following Equation 5.

[Equation 5]

$$f_{mot} = \Phi_{mot}(\Phi_{skl}(O_n))$$

According to the exemplary embodiments, a caption 242 describing the motion of the object may be extracted in a case where the information 202 associated with the object is input into a caption extractor 240. Here, the caption extractor 240 may include a video captioning model such as a SwinBERT model. In addition, the text feature vector f_{text} 252 may be extracted in a case where the caption 242 is input into a text extractor 250. Here, the text extractor 250 may include a model such as SimCSE for extracting sentence-based text features on the basis of the caption 242. For example, the text feature vector f_{text} 252 may be extracted as in the following Equation 6.

[Equation 6]

$$f_{text} = \Phi_{text}(\Phi_{cap}(O_n))$$

FIG. 3 is an exemplary view illustrating a structure of a diffusion model 300 according to the exemplary embodiments of the present disclosure. According to the exemplary embodiments, the diffusion model 300 may include an encoder 310 including a plurality of denoising attention blocks (DABs), a bottleneck (not shown), and a decoder 320. For example, the diffusion model 300 may generate a restoration vector 322 by removing noise injected into a noise vector 312 through the illustrated structure in FIG. 3.

According to the exemplary embodiments, the diffusion model 300 may generate the restoration vector 322 obtained by restoring the visual feature vector by taking the noise vector 312 as input and referencing the text feature vector or the motion feature vector as a condition 314.

That is, when restoring the noise vector 312, the diffusion model 300 may generate the restoration vector 322 to be close to the right answer by referring to the text or motion corresponding to the object.

According to the exemplary embodiments, in a case where the noise vector 312 passes through the diffusion model 300, an amount of noise determined according to a time step may be removed. For example, in a case where the noise vector 312 passes through the diffusion model 300 once, the amount of noise corresponding to one time step may be removed. In another example, in a case where the noise vector 312 passes through the diffusion model 300 once, an amount of noise corresponding to half the time step may also be removed. That is, in a case where the noise vector 312 passes through the diffusion model 300 iteratively for as long as an interval of a time step given in a noise injection process, the restoration vector 322 with all the noise removed may be generated.

FIG. 4 is an exemplary view illustrating a structure of a denoising attention block according to the exemplary embodiments of the present disclosure. According to the exemplary embodiments, the denoising attention block may be composed of a stack of a transformer block 410 and a residual block 420. In addition, the transformer block 410 may include a self-attention layer 416, a cross-attention layer 414, and a feed forward network 412. The residual block 420 may include a plurality of linear layers 422_1 and 422_2 connected by skip connection.

According to the exemplary embodiments, since the diffusion model operates depending on a time step, a time embedding vector 404 may be combined with a condition 402, which is the text feature vector or the motion feature vector. For example, vector combination may be performed as in the following Equation 7.

[Equation 7]

$$f'_{cond} = W_1 f_{cond} + W_2 f_{ftime}$$

Here, f'_{cond} may indicate a combined vector, f_{cond} may indicate the condition 402 which

is the text feature vector or the motion feature vector, and f_{ftime} may indicate a time embedding vector 404. In addition, $W_1 \in \mathbb{R}^{D_{cond} \times D_{vis}}$ and $W_2 \in \mathbb{R}^{D_{time} \times D_{vis}}$ may indicate projection matrices to respectively match a dimension D_{cond} of the condition 402 and a dimension D_{time} of the time embedding vector 404 with a dimension D_{vis} of an input visual feature vector 406. That is, the combined vector generated in this way may be provided to both of the transformer block 410 and the residual block 420 to help the diffusion model refer to the condition 402 more effectively.

According to the exemplary embodiments, the transformer block 410 may be a block for recognizing correlations between visual features and conditional features. To recognize these correlations, a segment-level multimodal feature vector may be converted into a clip-level multimodal feature vector. Here, a clip may be composed of a set of eight consecutive segments, but is not limited thereto. In this case, the clip-level multimodal feature vector may have a form $\mathbb{R}^{C \times 8 \times d}$ $C = S/16$. For example, the transformer block 410 may perform a calculation as in the following Equation 8.

[Equation 8]

$$TA(Q, K, V) = Softmax(\frac{QK^T}{\sqrt{d}})V$$

Here, it may be that $Q = W_q f_{vis} \in \mathbb{R}^{C \times 8 \times d}$, $K = W_k f'_{cond} \in \mathbb{R}^{C \times 8 \times d}$, $V = W_v f'_{cond} \in \mathbb{R}^{C \times 8 \times d}$. In addition, $W_q, W_k, W_v \mathbb{R}^{D_{vis} \times d}$ may indicate respective projection matrices, and d may indicate a dimension of query, key, and value. As described above, by the text feature vector and/or motion feature vector referenced as the conditions 402 when the transformer block 410 and the residual block 420 are calculated, the computing device 100 in FIG. 1 may effectively perform noise removal and vector restoration by referring to both of the text for describing the object and/or the motion of the object together with the visual features of the object.

FIG. 5 is a view illustrating an example in which a restoration process is performed by a

first diffusion model 520 and a second diffusion model 530 according to the exemplary embodiments of the present disclosure. As described above, the computing device 100 in FIG. 1 may input a noise vector into a diffusion model and generate a restoration vector with removed noise by using a text feature vector 512 and a motion feature vector 514 as conditions. Here, the diffusion model may include the first diffusion model 520 and the second diffusion model 530.

According to the exemplary embodiments, a first noise vector f_{vis}^{τ} 504 may be generated by a diffusion process 510 that injects noise into a visual feature vector f_{vis}^{0} 502. For example, in a case where a time step is set to τ, an amount of Gaussian noise corresponding to the time step τ may be injected into the visual feature vector f_{vis}^{0} 502 to generate the first noise vector f_{vis}^{τ} 504.

According to the exemplary embodiments, the first noise vector f_{vis}^{τ} 504 generated in this way may be input to the first diffusion model 520 trained with the text feature vector 512 as a condition. In this case, the first diffusion model 520 may generate a second noise vector $\hat{f}_{vis}^{\tau-1}$ 506 by removing noise once with reference to the text feature vector 512. Then, the generated second noise vector $\hat{f}_{vis}^{\tau-1}$ 506 may be input to the second diffusion model 530 trained with the motion feature vector 514 as a condition. In this case, the second diffusion model 530 may generate a third noise vector $\hat{f}_{vis}^{\tau-2}$ 508 by removing noise once again with reference to the motion feature vector 514.

In this case, the generated third noise vector $\hat{f}_{vis}^{\tau-2}$ 508 is again provided to the first diffusion model 520, so that a cycle may be formed between the first diffusion model 520 and the second diffusion model 530. Through the above-described process, the restoration vector may be generated as a result that the noise is iteratively removed for as long as the time step τ. With such a configuration, both of the first diffusion model 520 and the second diffusion model 530 with respective conditions different from each other are used instead of using a single diffusion model, so that the restoration performance may be improved, whereby the video anomaly detection may be performed with higher accuracy.

FIG. 6 is a view illustrating an example in which a restoration process is performed by a first diffusion model 520 and a second diffusion model 530 according to a second exemplary embodiment of the present disclosure. Unlike the restoration process described above, the order and/or method of using the first diffusion model 520 and the second diffusion model 530 may be determined differently. In the example in FIG. 6, the amount of Gaussian noise corresponding to a time step τ is injected into a visual feature vector, whereby a first noise vector 602 may be generated. In this case, the first noise vector 602 may be iteratively input into the first diffusion model 520 for as long as the time step τ, and accordingly, the first diffusion model 520 may generate a first restoration vector 604 with all the noise removed.

Then, the diffusion process 510 for the first restoration vector 604 may be performed again. For example, the amount of Gaussian noise corresponding to the time step τ is injected into the first restoration vector 604 so that a second noise vector 606 may be generated. In this case, the second noise vector 606 may be iteratively input into the second diffusion model 530 for as long as the time step τ, and accordingly, the second diffusion model 530 may generate a second restoration vector 608 with all the noise removed.

FIG. 7 is a view illustrating an example in which a restoration process is performed by a first diffusion model 520 and a second diffusion model 530 according to a third exemplary embodiment of the present disclosure. Unlike the restoration process described above, the order and/or method of using the first diffusion model 520 and the second diffusion model 530 may be determined differently. In the example in FIG. 7, an amount of Gaussian noise corresponding to a time step τ may be injected into a visual feature vector, so as to generate a first noise vector 702. In this case, the first noise vector 602 may be provided to the first diffusion model 520, and accordingly, the first diffusion model 520 may generate a second noise vector 704 from which the noise has been removed once.

According to the exemplary embodiments, the diffusion process 510 that injects noise for

as long as one time step to the second noise vector 704 again may be performed, so as to generate a third noise vector 706. In this case, the third noise vector 706 may be provided to the second diffusion model 530, and accordingly, the second diffusion model 530 may generate a fourth noise vector 708 from which noise is removed once again. That is, the noise may be removed once each time the first diffusion model 520 and the second diffusion model 530 are cycled, and a restoration vector may be generated in a case of performing the corresponding cycle repeatedly for as long as the time step τ.

FIG. 8 is a view illustrating an example in which a restoration process is performed by a first diffusion model 520 and a second diffusion model 530 according to a fourth exemplary embodiment of the present disclosure. Unlike the restoration processes described above, the order and/or method of using the first diffusion model 520 and the second diffusion model 530 may be determined differently. In the example in FIG. 8, an amount of Gaussian noise corresponding to a time step τ may be injected into a visual feature vector to generate a first noise vector 802. In this case, the first noise vector 802 may be iteratively input into the first diffusion model 520 for as long as half the time step τ, and accordingly, the first diffusion model 520 may generate a second noise vector 804 with half the noise removed.

Then, the second noise vector 804 may be iteratively input into the second diffusion model 530 for as long as half the time step τ, and accordingly, the second diffusion model 530 may generate a restoration vector 806 with all the noise removed. As described above in FIGS. 5 to 8, the method of using the first diffusion model 520 and the second diffusion model 530 may be determined in various ways, and a restoration process may be performed to have optimal performance depending on the detection conditions of abnormalities and/or abnormal behaviors.

FIG. 9 is a flowchart illustrating an example of a method 900 for detecting a video anomaly on the basis of multimodal diffusion according to the exemplary embodiments of the present disclosure. The method 900 for detecting the video anomaly on basis of the multimodal

diffusion may be performed by a processor (e.g., at least one processor of the computing device). In step S910, the method 900 for detecting the video anomaly on basis of the multimodal diffusion may be initiated by the processor obtaining video data including a plurality of frames. For example, the video data may be obtained from a surveillance camera such as CCTV, but is not limited thereto.

According to the exemplary embodiments, in step S920, the processor may detect an object included in each of the plurality of frames. For example, the processor may detect the object included in each of the plurality of frames by using any object detection algorithm and/or machine learning model. Then, in step S930, the processor may extract a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object.

According to the exemplary embodiments, the processor may provide information related to the detected object to a trained I3D-based model, so as to extract the visual feature vector for the object. Additionally or alternatively, the processor may provide information related to the detected object to a BERT-based model to generate a caption describing the object, and provide the generated caption to the trained SimCSE-based model to extract the text feature vector corresponding to the description of the object. Additionally or alternatively, the processor may provide information related to the detected object to a trained HRNet-based model to extract skeletal information corresponding to the object and extract the motion feature vector representing the motion of the object by using the extracted skeletal information.

According to the exemplary embodiments, in step S940, the processor may generate a noise vector by injecting noise into the visual feature vector. For example, the processor may generate the noise vector by injecting an amount of Gaussian noise determined according to a range of a time step into the visual feature vector. Here, the noise vector refers to a vector in which at least some noise has been injected into the visual feature vector, and may include both a vector

initially generated by a diffusion process and a vector that has not sufficiently passed through a diffusion model and thus still has the remaining noise.

According to the exemplary embodiments, in step S950, the processor may input a noise vector into a diffusion model and generate a restoration vector with removed noise by using the text feature vector and the motion feature vector as conditions. Here, the diffusion model may include a first diffusion model and a second diffusion model. The processor may generate the restoration vector with the noise removed by iteratively performing: a first restoration step of inputting a noise vector into the first diffusion model and removing at least some noise included in the noise vector by using the text feature vector as a condition; and a second restoration step of inputting noise into the second diffusion model and removing at least some of the noise included in a noise vector by using the motion feature vector as a condition.

According to the exemplary embodiments, in step S960, the processor may perform anomaly detection on the video data by comparing the visual feature vector and the restoration vector. In this case, the processor may calculate an anomaly score based on a distance between the visual feature vector and the restoration vector, and perform the anomaly detection on the video data on the basis of whether the calculated anomaly score is greater than or equal to a threshold value. For example, the anomaly score may be calculated by using a mean squared error between the visual feature vector and the reconstruction vector.

FIG. 10 is a block diagram illustrating a hardware configuration of the computing device 100 according to the exemplary embodiments of the present disclosure. The computing device 100 may include a memory 1010, a processor 1020, a communication module 1030, and an input/output interface 1040. As shown in FIG. 10, the computing device 100 may be configured to communicate information and/or data through a network by using the communication module 1030.

The memory 1010 may include any non-transitory computer-readable recording medium.

According to the exemplary embodiments, the memory 1010 may include a permanent mass storage device such as a random access memory (RAM), a read only memory (ROM), a disk drive, a solid state drive (SSD), and a flash memory. As another example, the permanent mass storage device such as the ROM, SSD, flash memory, and disk drive may be included in the computing device 100 as a separate permanent storage device distinguished from the memory. In addition, an operating system and at least one program code may be stored in the memory 1010.

Such software components may be loaded from the computer-readable recording medium as a separate medium distinguished from the memory 1010. Such separate computer-readable recording medium may include a recording medium directly connectable to such a computing device 100, and may include, for example, the computer-readable recording medium such as a floppy drive, a disk, a tape, a DVD/CD-ROM drive, and a memory card. As another example, the software components may be loaded into the memory 1010 through the communication module 1030 other than the computer-readable recording medium. For example, at least one program may be loaded into the memory 1010 on the basis of a computer program installed through a file provided by developers or a file distribution system distributing an installation file of an application through the communication module 1030.

The processor 1020 may be configured to process commands of the computer program by performing fundamental arithmetic, logic, and input/output calculations. The commands may be provided to another user terminal (not shown) or another external system by the memory 1010 or the communication module 1030.

The communication module 1030 may provide a component or function for the user terminal (not shown) and the computing device 100 to communicate with each other through a network, and may provide a component or function for the computing device 100 to communicate with an external system (e.g., a separate cloud system, etc.). For example, control signals, commands, data, and the like provided under the control of the processor 1020 of the computing

device 100 may be transmitted to the user terminal and/or the external system through communication modules of the user terminal and/or the external system via the communication module 1030 and the network.

In addition, the input/output interface 1040 of the computing device 100 may be a means for interfacing with a device (not shown) for input or output, the device being connectable to the computing device 100 or included by the computing device 100. In FIG. 10, the input/output interface 1040 is illustrated as a component configured separately from the processor 1020, but is not limited thereto, and the input/output interface 1040 may be configured to be included in the processor 1020. The computing device 100 may include more components than those in FIG. 10. However, there is no need to explicitly illustrate most of the components of the related art.

The processor 1020 of the computing device 100 may be configured to manage, process, and/or store information and/or data received from a plurality of user terminals and/or a plurality of external systems.

The above-described method and/or various exemplary embodiments may be realized by digital electronic circuits, computer hardware, firmware, software, and/or a combination thereof. The various exemplary embodiments of the present disclosure may be executed by data processing devices, for example, one or more programmable processors and/or one or more computing devices, or may be implemented with computer-readable recording media and/or computer programs stored on the computer-readable recording media. The computer programs described above may be written in any types of programming languages, including compiled or interpreted languages, and may be distributed in any types thereof, such as standalone programs, modules, and subroutines. The computer programs may be distributed through a single computing device, a plurality of computing devices connected through the same network, and/or a plurality of computing devices distributed so as to be connected through a plurality of different networks.

The above-described method and/or various exemplary embodiments may be performed

by one or more processors configured to execute one or more computer programs that process, store, and/or manage any feature, function, and the like by operating on the basis of input data or generating output data. For example, the method and/or various exemplary embodiments of the present disclosure may be performed by a special purpose logic circuit such as a Field Programmable Gate Array (FPGA) or an Application Specific Integrated Circuit (ASIC), and the device and/or the system for performing the method and/or exemplary embodiments of the present disclosure may be implemented by using the special purpose logic circuits such as the FPGA or ASIC.

The one or more processors for executing the computer programs may include general purpose or special purpose microprocessors and/or one or more processors of any type of digital computing device. Each processor may receive commands and/or data from each of the read-only memory and random access memory, or may receive commands and/or data from all the read-only memory and random access memory. In the present disclosure, the components of the computing device for performing the method and/or exemplary embodiments may include one or more processors for executing the commands, and one or more memory devices for storing the commands and/or data.

According to the exemplary embodiments, the computing device may exchange data with one or more mass storage devices for storing data. For example, the computing device may receive data from a magnetic disc or an optical disc, and may transmit data to the magnetic disc or the optical disc. The computer-readable storage medium suitable for storing the commands and/or data associated with the computer programs may include, but is not limited to, any type of non-volatile memory including semiconductor memory devices such as Erasable Programmable Read-Only Memory (EPROM), Electrically Erasable PROM (EEPROM), and flash memory devices. For example, the computer-readable storage medium may include: the magnetic disk such as an internal hard disk or a removable disk; a magneto-optical disk; a CD-ROM; and a DVD-ROM

disk.

To provide interaction with a user, a computing device may include, but is not limited to, a display device (e.g., a Cathode Ray Tube (CRT), a Liquid Crystal Display (LCD), etc.) for providing or displaying information to the user, and a pointing device (e.g., a keyboard, a mouse, a trackball, etc.) for allowing the user to provide input and/or commands, and the like to the computing device. That is, the computing device may further include any other type of device for providing the interaction with the user. For example, for the interaction with the user, the computing device may provide any form of sensory feedback including visual feedback, auditory feedback, and/or tactile feedback to the user. In this regard, the user may provide the input to the computing device through various gestures of vision, voice, motion, etc.

In the present disclosure, the various exemplary embodiments may be implemented in a computing system including a backend component (e.g., a data server), a middleware component (e.g., an application server), and/or a frontend component. In this case, the components may be interconnected by any form or medium of digital data communication, such as a communication network. For example, the communication network may include a Local Area Network (LAN), a Wide Area Network (WAN), etc.

The computing device based on the exemplary embodiments described in the present specification may be implemented by using hardware and/or software configured to interact with a user, including a user device, a user interface (UI) device, a user terminal, or a client device. For example, the computing device may include a portable computing device such as a laptop computer. Additionally or alternatively, the computing device may include, but is not limited to, a personal digital assistant (PDA), a tablet PC, a game console, a wearable device, an internet of things (IoT) device, a virtual reality (VR) device, an augmented reality (AR) device, etc. The computing device may further include other types of devices configured to interact with the user. In addition, the computing device may include a portable communication device (e.g., a mobile

phone, a smartphone, a wireless cellular phone, etc.) suitable for wireless communication over a network, such as a mobile communication network. The computing device may be configured to communicate wirelessly with a network server by using wireless communication technologies and/or protocols, such as Radio Frequency (RF), Microwave Frequency (MWF), and/or Infrared Ray Frequency (IRF).

The various exemplary embodiments including specific structural and functional details in the present disclosure are exemplary. Therefore, the exemplary embodiments of the present disclosure are not limited to those described above and may be implemented in various other forms. In addition, the terminology used in the present disclosure is for the purpose of describing some exemplary embodiments, and is not to be construed as limiting the exemplary embodiments. For example, words and the terms in singular form described above may be interpreted as to include those in plural form as well, unless the context clearly indicates otherwise.

In the present disclosure, unless otherwise defined, all terms used in the present specification, including technical or scientific terms, have the same meaning as commonly understood by those skilled in the art to which such concepts belong. In addition, commonly used terms, such as terms defined in dictionaries, should be interpreted to have a meaning consistent with their meaning in the context of the related art.

Although the present disclosure has been described in relation to some exemplary embodiments in the present specification, various modifications and changes may be made without departing from the scope of the embodiments of the present disclosure that may be understood by those skilled in the art to which the embodiments of the present disclosure pertains. Furthermore, such modifications and changes should be considered to fall within the scope of the claims appended to the present specification.

WHAT IS CLAIMED IS:

1. A method of detecting a video anomaly and being performed by at least one processor, the method comprising:

obtaining video data comprising a plurality of frames;

detecting an object included in each of the plurality of frames;

extracting a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object;

generating a noise vector by injecting noise into the visual feature vector;

generating a restoration vector with the noise removed by inputting the noise vector into a diffusion model and using the text feature vector and the motion feature vector as conditions; and

performing anomaly detection on the video data by comparing the visual feature vector and the restoration vector.

2. The method of claim 1, wherein the extracting of the multimodal feature vector comprises:

extracting the visual feature vector for the object by providing information related to the detected object to a trained model based on Inflated 3D ConvNet (I3D).

3. The method of claim 1, wherein the extracting of the multimodal feature vector comprises:

generating a caption for describing the object by providing information related to the detected object to a model based on Bidirectional Encoder Representations from Transformers (BERT); and

extracting the text feature vector corresponding to the description of the object by

providing the generated caption to a trained model based on Simple Contrastive Learning of Sentence Embeddings (SimCSE).

4. The method of claim 1, wherein the extracting of the multimodal feature vector comprises:

extracting skeletal information corresponding to the object by providing information related to the detected object to a trained model based on High-Resolution Network (HRNet); and

extracting the motion feature vector representing motion of the object by using the extracted skeletal information.

5. The method of claim 4, wherein the extracting of the motion feature vector representing the motion of the object by using the extracted skeletal information comprises:

extracting the motion feature vector by providing the extracted skeletal information to a trained model based on PoseConv3D.

6. The method of claim 1, wherein the generating of the noise vector by injecting the noise into the visual feature vector comprises:

generating the noise vector by injecting an amount of Gaussian noise determined according to a range of a time step into the visual feature vector.

7. The method of claim 1, wherein the diffusion model includes a first diffusion model and a second diffusion model, and

the generating of the restoration vector with the noise removed comprises:

a first restoration step of inputting the noise vector into the first diffusion model and removing at least some of the noise included in the noise vector by using the text feature vector as

a condition; and

a second restoration step of inputting a noise vector into the second diffusion model and removing at least some of the noise included in the noise vector by using the motion feature vector as a condition.

8. The method of claim 7, wherein the generating of the restoration vector with the noise removed further comprises:

generating the restoration vector with the noise removed by iteratively performing the first restoration step and the second restoration step.

9. The method of claim 1, wherein the performing of the anomaly detection on the video data comprises:

calculating an anomaly score based on a distance between the visual feature vector and the restoration vector; and

performing the anomaly detection on the video data on the basis of whether the calculated anomaly score is greater than or equal to a threshold value.

10. The method of claim 9, wherein the calculating of the anomaly score comprises:

calculating the anomaly score according to the distance by using a mean square error (MSE) between the visual feature vector and the restoration vector.

11. The method of claim 1, wherein the diffusion model comprises:

an encoder comprising a plurality of denoising attention blocks (DABs);

a bottleneck; and

a decoder.

12. The method of claim 11, wherein each denoising attention block comprises:

a residual block comprising a plurality of linear layers connected by skip connection; and

a transformer block comprising a self-attention layer, a cross-attention layer, and a feed-forward network (FFN).

13. A non-transitory computer readable recording medium storing computer program to execute a method of detecting a video anomaly on a computer according to claim 1.

14. A computing device comprising:

a communication module;

a memory; and

at least on processor connected to the memory and configured to execute at least one computer-readable program comprised in the memory,

wherein the at least one program comprises:

commands that obtain video data including a plurality of frames, detect an object included in each of the plurality of frames, extract a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object, generate a noise vector by injecting noise into the visual feature vector, generate a restoration vector with the noise removed by inputting the noise vector into a diffusion model and using the text feature vector and the motion feature vector as conditions, and perform anomaly detection on the video data by comparing the visual feature vector and the restoration vector.

ABSTRACT

Proposed are a method of detecting a video anomaly on the basis of multimodal diffusion, and the method includes a step of obtaining video data including a plurality of frames, a step of detecting an object included in each of the plurality of frames, a step of extracting a multimodal feature vector including a visual feature vector, a text feature vector, and a motion feature vector for the detected object, a step of generating a noise vector by injecting noise into the visual feature vector, a step of generating a restoration vector with the noise removed by inputting the noise vector into a diffusion model and by using the text feature vector and the motion feature vector as conditions, and a step of performing anomaly detection on the video data by comparing the visual feature vector and the restoration vector.

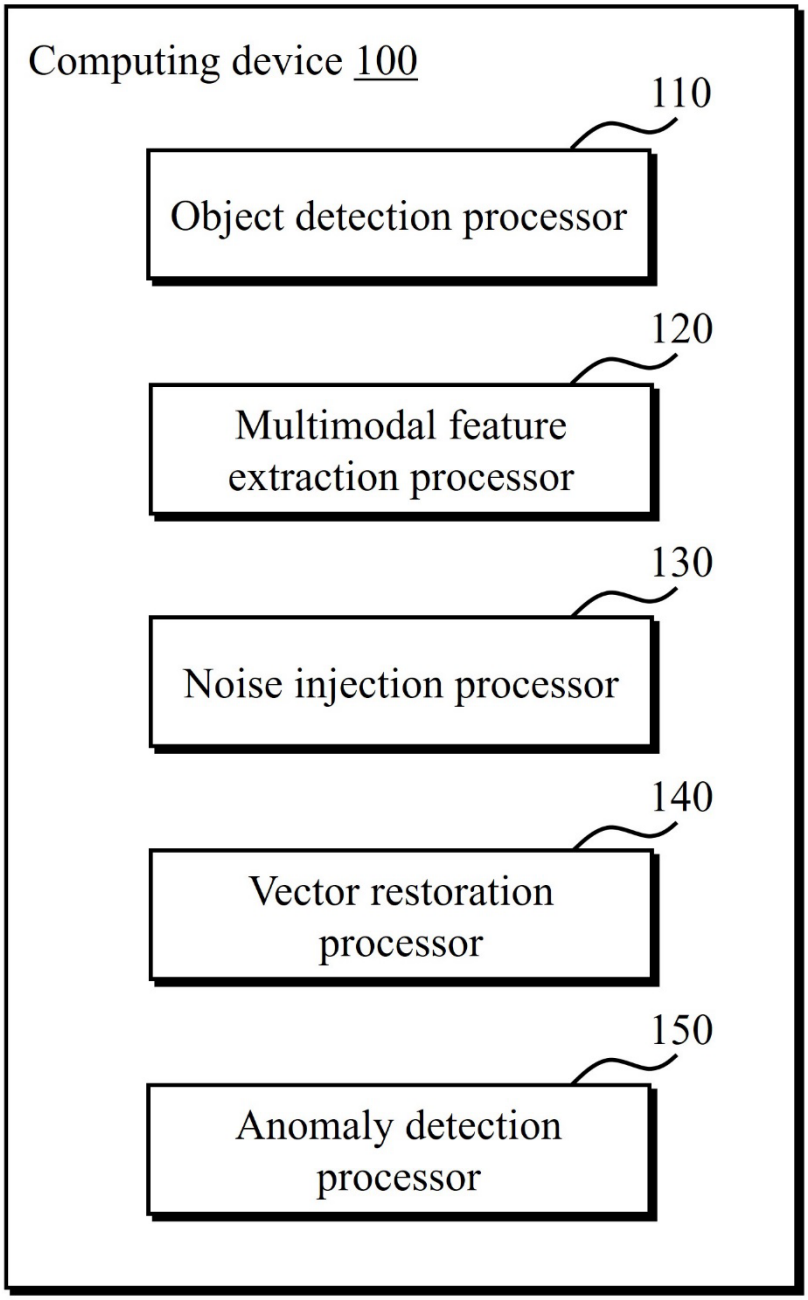
Computing device 100
110
Object detection processor
120
Multimodal feature extraction processor
130
Noise injection processor
140
Vector restoration processor
150
Anomaly detection processor

FIG. 1

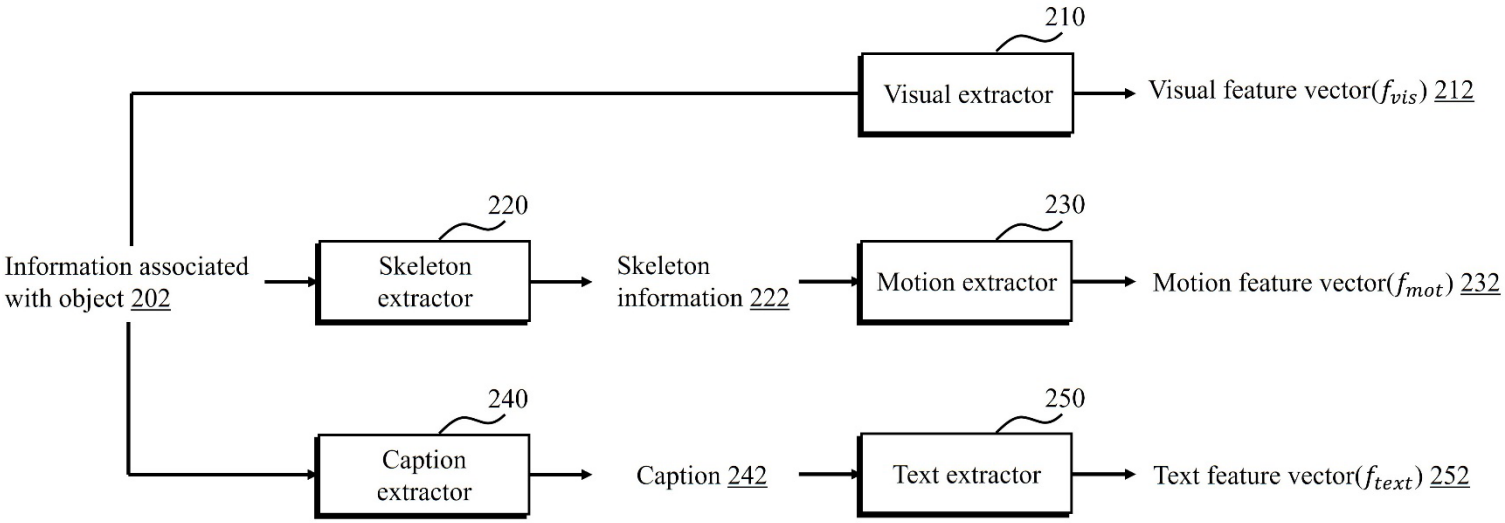
210
Visual extractor
Visual feature vector(f_{vis}) 212
Information associated with object 202
220
Skeleton extractor
Skeleton information 222
230
Motion extractor
Motion feature vector(f_{mot}) 232
240
Caption extractor
Caption 242
250
Text extractor
Text feature vector(f_{text}) 252

FIG. 2

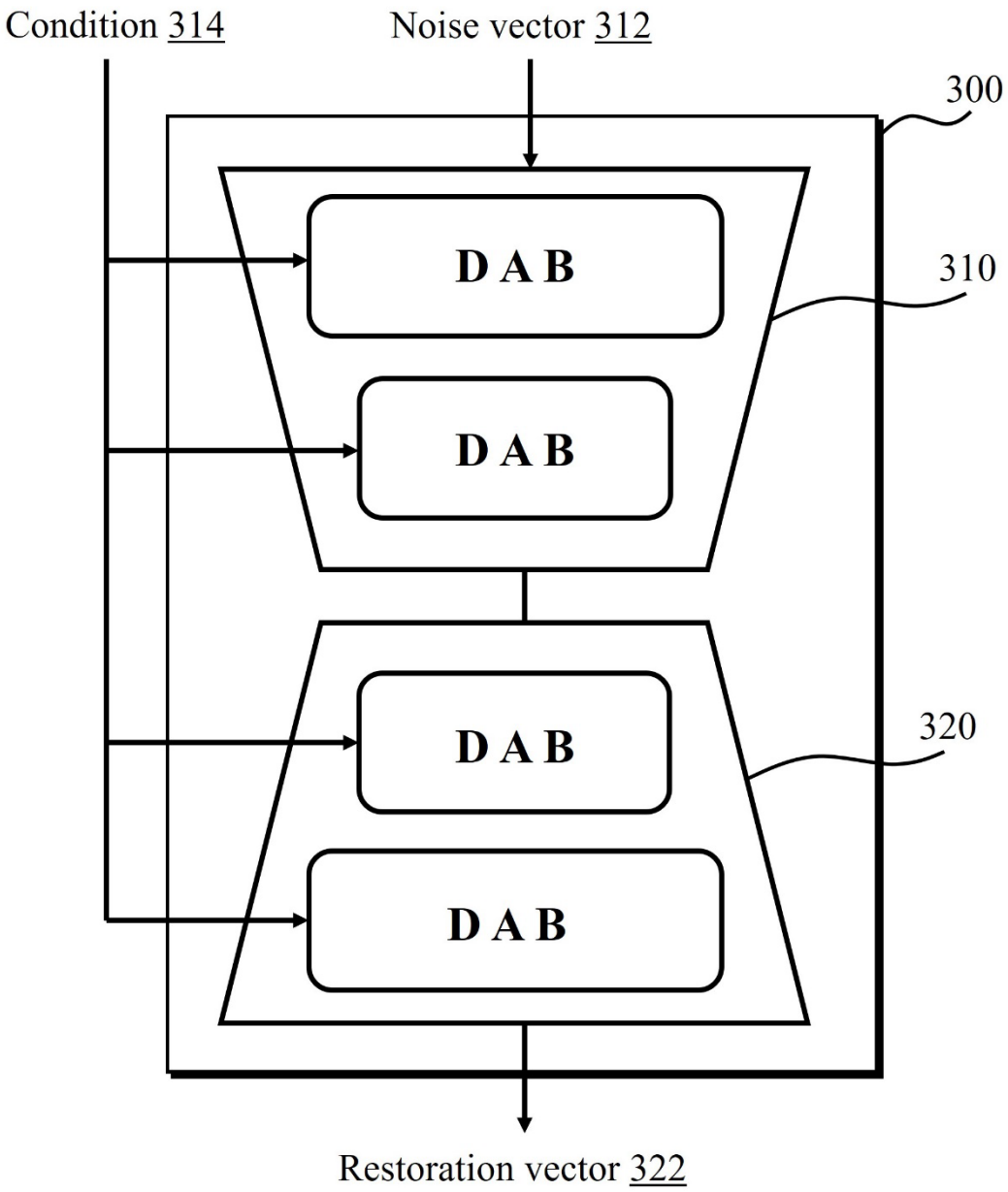
Condition 314
Noise vector 312
300
D A B
D A B
310
D A B
D A B
320
Restoration vector 322

FIG. 3

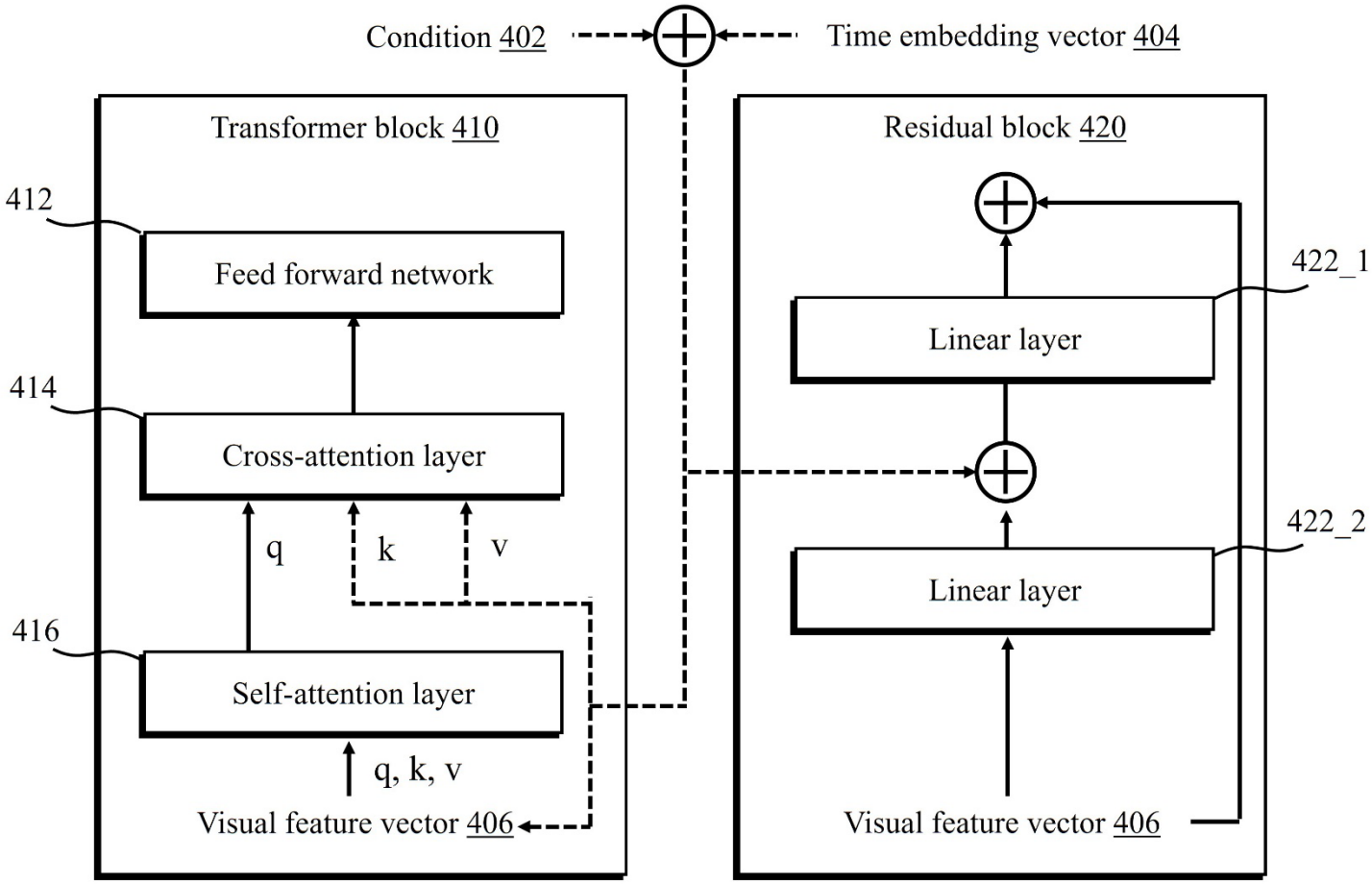
Condition 402
Time embedding vector 404
Transformer block 410
Residual block 420
412
Feed forward network
414
Cross-attention layer
q
k
v
416
Self-attention layer
q, k, v
Visual feature vector 406
422_1
Linear layer
422_2
Linear layer
Visual feature vector 406

FIG. 4

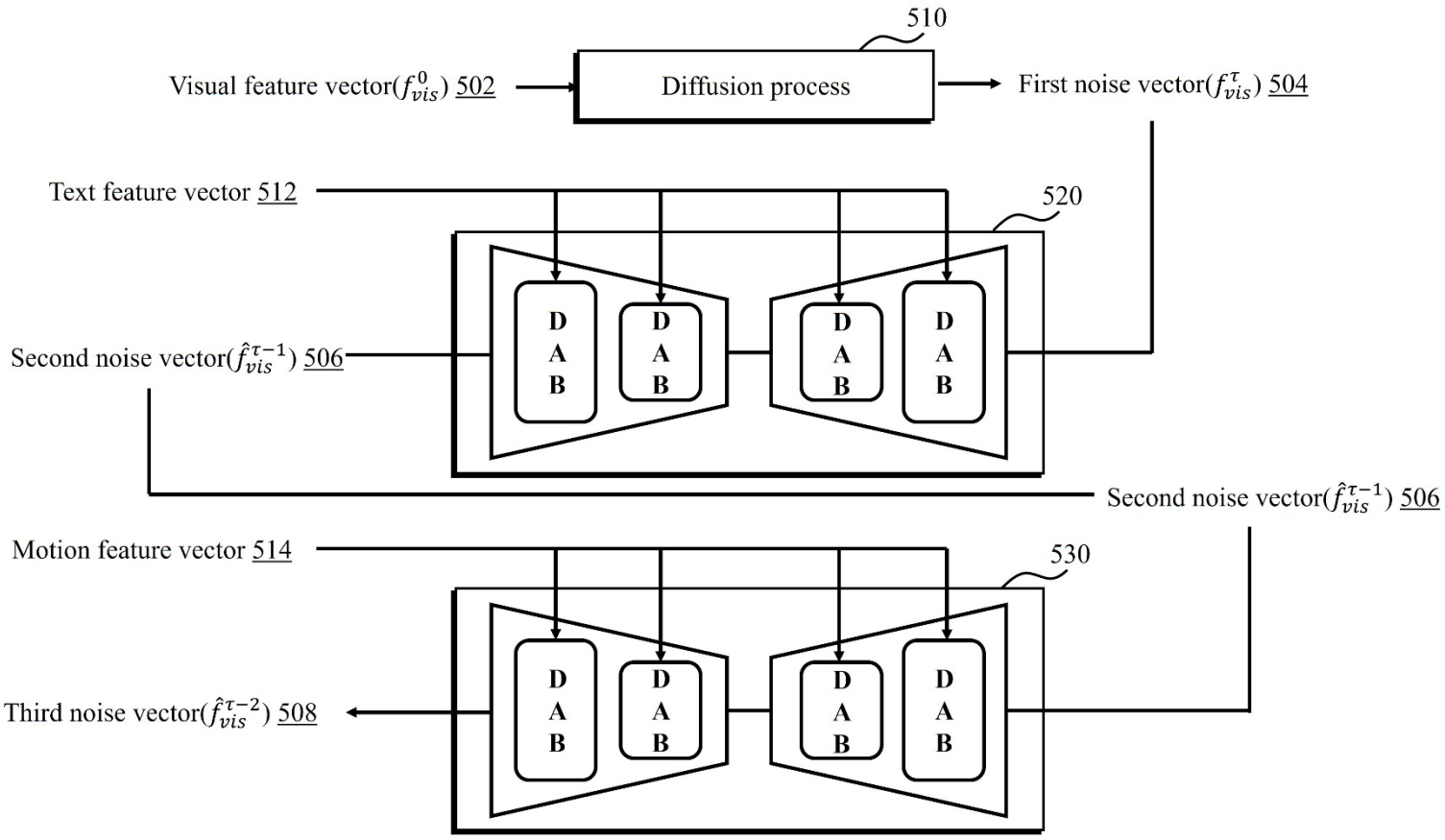
510
Visual feature vector(f_{vis}^{0}) 502
Diffusion process
First noise vector(f_{vis}^{τ}) 504
Text feature vector 512
520
D A B
D A B
D A B
D A B
Second noise vector($\hat{f}_{vis}^{\tau-1}$) 506
Second noise vector($\hat{f}_{vis}^{\tau-1}$) 506
Motion feature vector 514
530
D A B
D A B
D A B
D A B
Third noise vector($\hat{f}_{vis}^{\tau-2}$) 508

FIG. 5

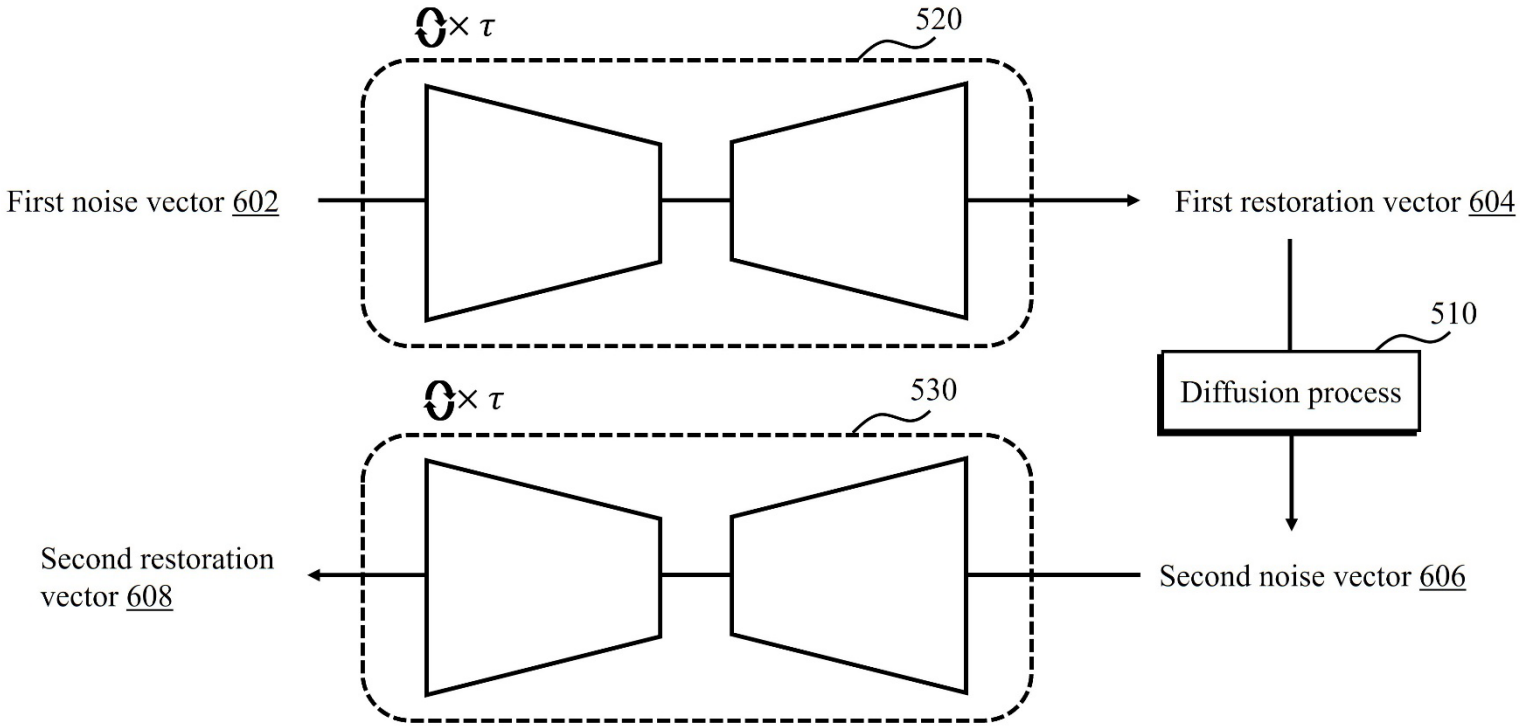
× τ
520
First noise vector 602
First restoration vector 604
510
Diffusion process
× τ
530
Second restoration vector 608
Second noise vector 606

FIG. 6

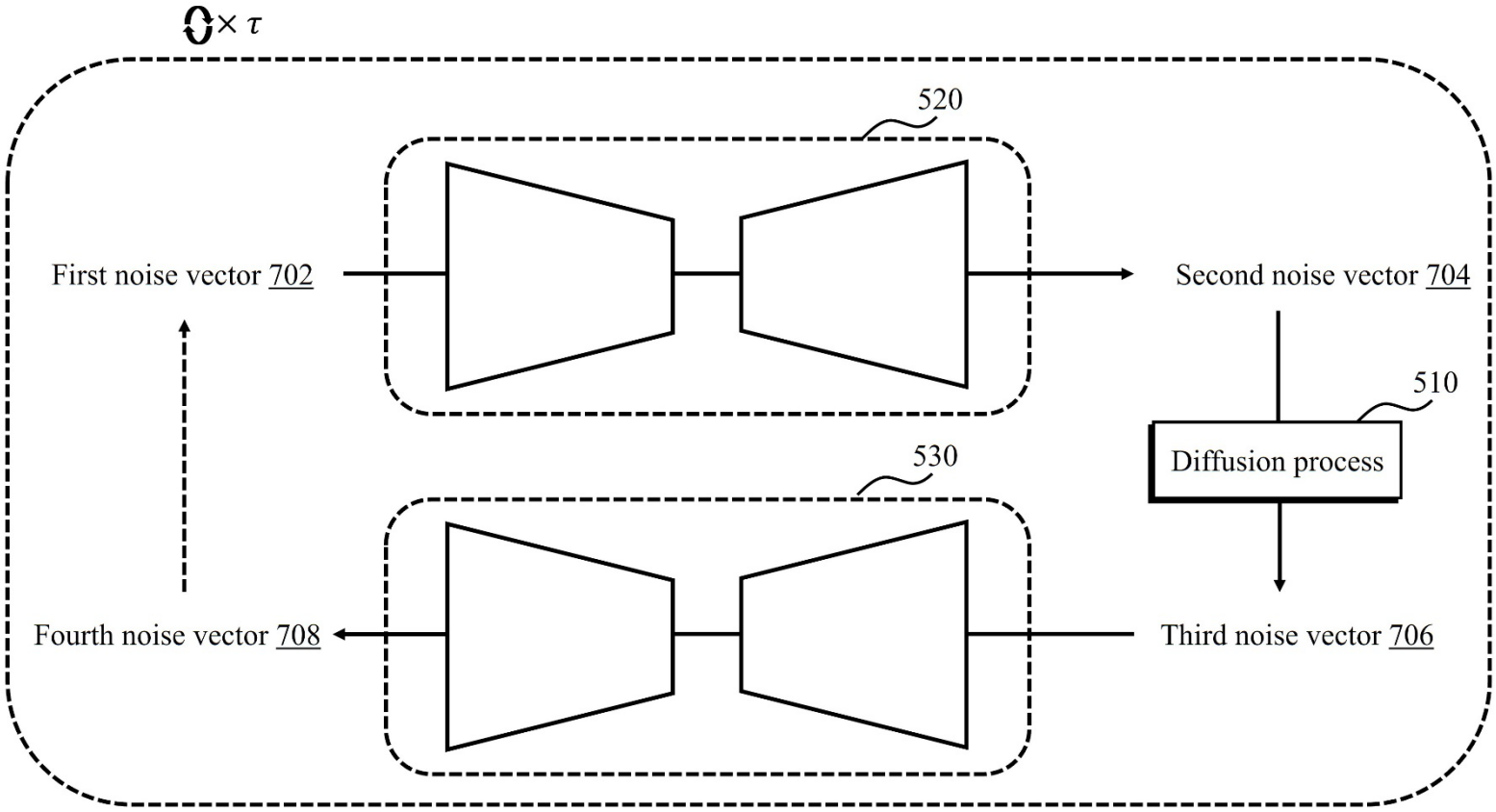
0×τ
520
First noise vector 702
Second noise vector 704
510
Diffusion process
530
Fourth noise vector 708
Third noise vector 706

FIG. 7

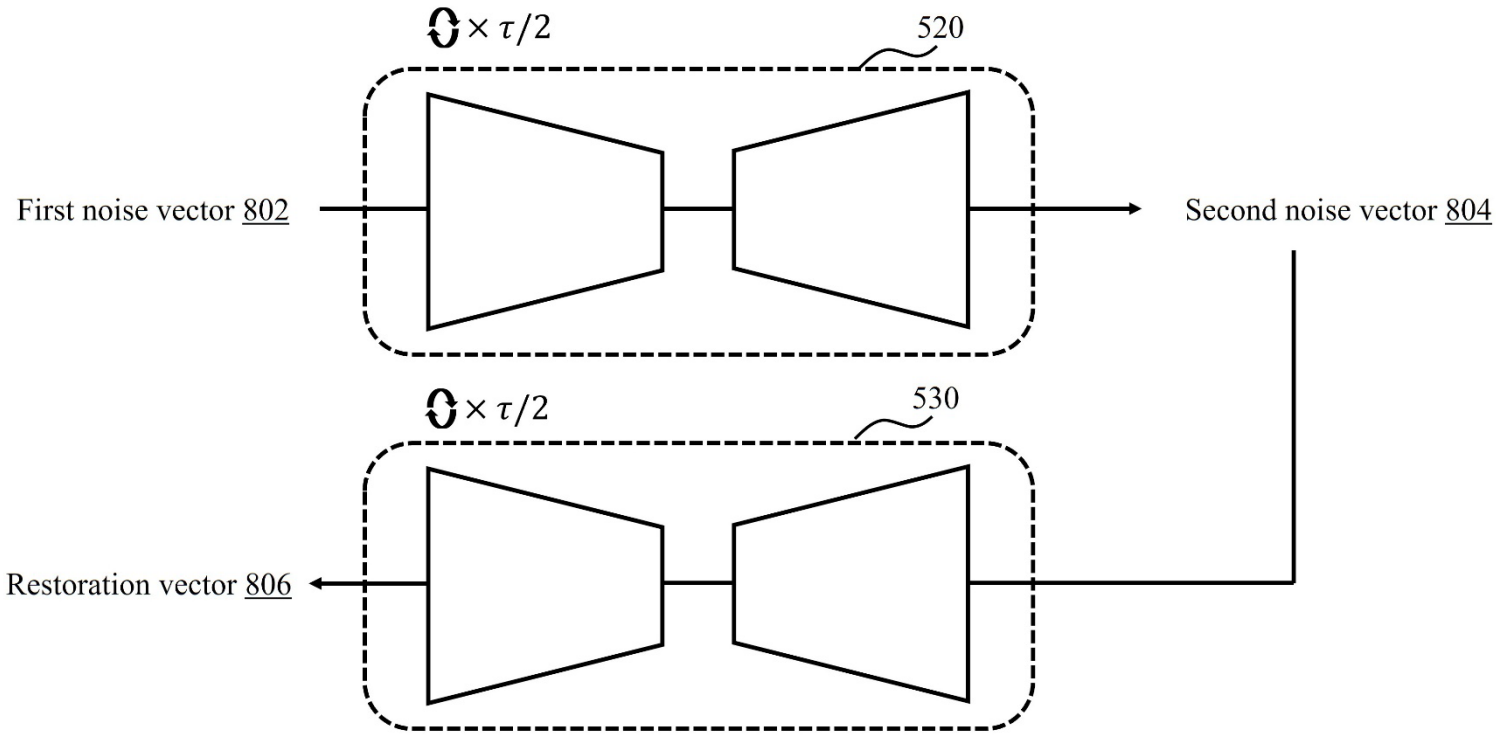
× τ/2
520
First noise vector 802
Second noise vector 804
× τ/2
530
Restoration vector 806

FIG. 8

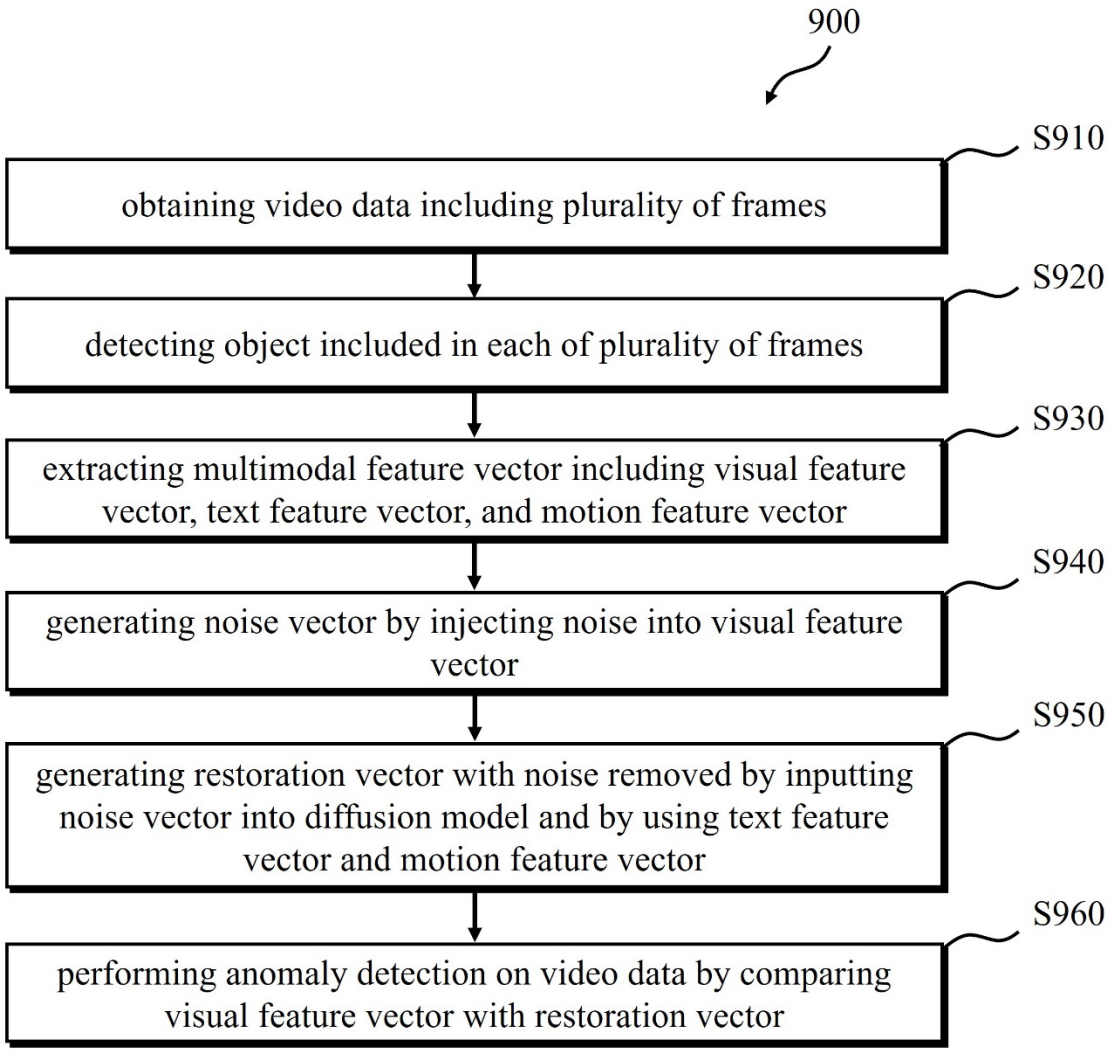
900
S910
obtaining video data including plurality of frames
S920
detecting object included in each of plurality of frames
S930
extracting multimodal feature vector including visual feature vector, text feature vector, and motion feature vector
S940
generating noise vector by injecting noise into visual feature vector
S950
generating restoration vector with noise removed by inputting noise vector into diffusion model and by using text feature vector and motion feature vector
S960
performing anomaly detection on video data by comparing visual feature vector with restoration vector

FIG. 9

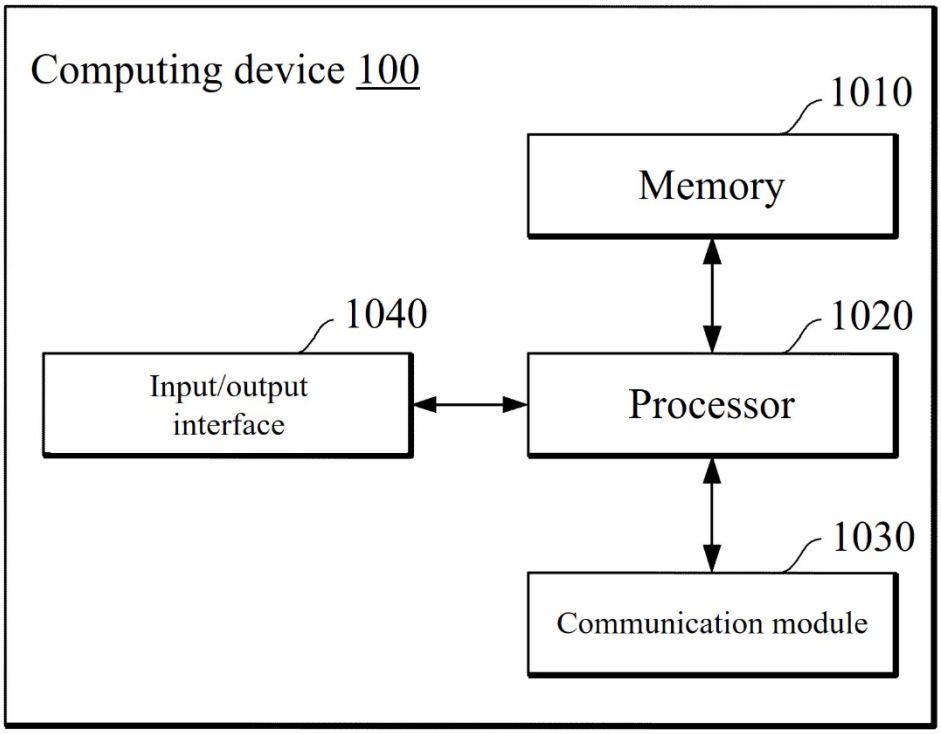
Computing device 100
1010
Memory
1040
Input/output interface
1020
Processor
1030
Communication module

FIG. 10

COMBINED DECLARATION (37 CFR §1.63) FOR UTILITY PATENT APPLICATION AND ASSIGNMENT FORM[1]

The undersigned acknowledges that this document is being used both as an assignment of the invention and as the declaration (37 CFR 1.63) for a Utility or Design Application.

WHEREAS, the undersigned, hereinafter referred to collectively as Assignor, has invented:

Title: **METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR**

for which Assignor is about to make or has made United States or International application for patent.

WHEREAS, as a below named inventor(s), I/(we) hereby declare that:

Section I. Declaration

This declaration is directed to:

☒ The attached U.S. non-provisional patent application, or

☐ U.S. non-provisional patent application number _____, filed on _____ , or

☐ PCT international patent application number _____, filed on _____ .

The above-identified application was made or authorized to be made by me (us).

I (we) believe that I am (we are) the original inventor (original joint inventors) of a claimed invention in the above-identified application.

I (we) hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

I hereby state that I have reviewed and understand the contents of the above identified application, including the claims.

I acknowledge that I am aware of the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56.

[1] This form requires the use of an Application Data Sheet.

Section II. Assignment

In consideration of the sum of One Dollar ($1.00) and other good and valuable consideration paid to each of the undersigned, the undersigned hereby sell(s) and assign(s) to

UIF (University Industry Foundation), Yonsei University

having an address at 50 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea (hereinafter designated as the Assignee), the entire (100%) right, title and interest for the United States as defined in 35 USC §100, in the invention described in the application identified in Section I of this document.

Assignor hereby confirms any prior assignment to Assignee, and to the extent that Assignor has not already done so, agrees to assign, and hereby does, sell, assign and transfer unto Assignee and its successors in interest, the full and exclusive right, title and interest in the United States of America and throughout the world, including the right to claim priority under the laws of the United States, the Paris Convention, and any foreign countries, to the inventions as described in the aforesaid application, to the aforesaid application itself, and all divisions, continuations, continuations-in-part, or other applications claiming priority directly or indirectly from the aforesaid application, and any United States or foreign Letters Patent, utility model, or other similar rights which may be granted thereon, including reissues, reexaminations and extensions thereof, and all copyright rights throughout the world in the aforesaid application and the subject matter disclosed therein, these rights, title and interest to be held and enjoyed by Assignee to the full end of the term for which the Letters Patent, utility model, or other similar rights, are granted and any extensions thereof as fully and entirely as the same would have been held by Assignor had this assignment and sale not been made, and the right to sue for, and recover for past infringements of, or liabilities for, any of the rights relating to any of the applications, patents, utility models, or other similar rights, resulting therefrom, and the copyright rights;

Assignor hereby covenants and agrees to execute all instruments or documents required or requested for the making and prosecution of any applications of any type for patent, utility model, or other similar rights, and for copyright, in the United States and in all foreign countries including, but not limited to, any provisional, continuation, continuation-in-part, divisional, renewal or substitute thereof, and as to letters patent any reissue, re-examination, or extension thereof, and for litigation regarding, or for the purpose of protecting title and to the said invention, the United States application for patent, or Letters Patent therefor, and to testify in support thereof, for the benefit of Assignee without further or other compensation than that above set forth;

Assignor hereby covenants that no assignment, sale, license, agreement or encumbrance has been or will be entered into which would conflict with this Assignment; and

Assignor hereby requests the United States Patent and Trademark Office to issue the Letters Patent of the United States of America to Assignee, and requests that any official of any country or countries foreign to the United States, whose duty it is to issue or grant patents and

applications as aforesaid, to issue the Letters Patent, Utility Model Registration or other similar right to Assignee.

The undersigned hereby grant(s) the law firm of **ZION IP** the power to insert on this Declaration and Assignment any further identification which may be necessary or desirable in order to comply with the rules of the U.S. Patent and Trademark Office for recordation of this document.

Section III. Inventor(s)/Assignor(s) Signature(s)

LEGAL NAME OF FIRST INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	PARK	Sanghyun	
SIGNATURE 		DATE 25.03.12	

LEGAL NAME OF SECOND INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	LEE	Kijung	
SIGNATURE 이기정		DATE 25.03.12	

LEGAL NAME OF THIRD INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	JO	Youngwan	
SIGNATURE 조영완		DATE 25.03.12	

LEGAL NAME OF FOURTH INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	AHN	Sunghyun	
SIGNATURE  DATE 25.03.12			

UNITED STATES PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

ELECTRONIC ACKNOWLEDGEMENT RECEIPT

APPLICATION #	RECEIPT DATE / TIME	ATTORNEY DOCKET #
19/078,589	**03/13/2025 12:33:59 PM Z ET**	**PP6807264**

Title of Invention

METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR

Application Information

APPLICATION TYPE	Utility - Nonprovisional Application under 35 USC 111(a)	PATENT #	-
CONFIRMATION #	4817	FILED BY	BYUNGWOONG PARK
PATENT CENTER #	69565003	FILING DATE	-
CUSTOMER #	183715	FIRST NAMED INVENTOR	Sanghyun PARK
CORRESPONDENCE ADDRESS	-	AUTHORIZED BY	-

Documents

TOTAL DOCUMENTS: 16

DOCUMENT	PAGES	DESCRIPTION	SIZE (KB)
2-PP6807264-ADS.pdf	9	Application Data Sheet	1227 KB
7-PP6807264 Specifications-APP.TEXT.docx	32	Application body structured text document	82 KB
Warning: Bookmarks were found and have been removed.			
6-5-Deep Video Anomaly Detection Opportunities and Challenges.pdf	8	Non Patent Literature	500 KB
1-PP6807264-Transmittal.pdf	2	Transmittal of New Application	400 KB
6-3-KR10-2323671-TR.pdf	28	Foreign Reference	1730 KB
3-PP6807264-POA.pdf	2	Power of Attorney	181 KB

4-PP6807264-Authorization for internet communication.pdf	2	Internet Communications Authorization	182 KB
5-PP6807264-PDX.pdf	2	Request for USPTO to retrieve priority docs	306 KB
6-0-PP6807264-IDS.pdf	4	Information Disclosure Statement (IDS) Form (SB08)	1010 KB
8-PP6807264-Drawing.pdf	10	Drawings-only black and white line drawings	3438 KB
6-4-MDVAD Multimodal Diffusion for Video Anomaly Detection.pdf	18	Non Patent Literature	5296 KB
6-1-KR10-2022-0088653-TR.pdf	34	Foreign Reference	8575 KB
9-PP6807264_Declaration_Assignment__.pdf	4	Oath or Declaration filed	1236 KB
6-2-JP2018-077479-TR____part_1.pdf	100	Foreign Reference	13567 KB
6-2-JP2018-077479-TR____part_2.pdf	98	Foreign Reference	14267 KB
7-PP6807264 Specifications.pdf	32	Auxiliary PDF of Application	226 KB

Digest

DOCUMENT	MESSAGE DIGEST(SHA-512)
2-PP6807264-ADS.pdf	5A32FBA9155BA15E660F37DBC4D0BB8EC903C799B4C21331551AB3551F9601C3288950DAF9499149B13B206FC9015644BD267BFB63AC1EF018717FB7070501BB
7-PP6807264 Specifications-APP.TEXT.docx	9BB6047239A0B34047774655BC4BAA1A7B74FC82C94555F230CBCE3DEB7CEDC2B2200846291E25E5D9A233972AF47330C564BBD4F141C2859C681F3263DAEDAE
6-5-Deep Video Anomaly Detection Opportunities and Challenges.pdf	36BC9E42D707B16CBED2AB50921BEC32DD9A1846D19AF62B8675D679BBDD176E6173953F210FC6C1D4788AB53AF3518758D8D46E73FC08F948E9A1974E6BB86C
1-PP6807264-Transmittal.pdf	FFA0867BC128BD83C1BB1A16C914C0B38175CFDB4D04DAF30E58970F1A0AA0EF7D111DD904450A86266574A425828E8C2CD0BB32A352ADCAE384F8C6940234DD

6-3-KR10-2323671-TR.pdf	6C35BE9363AB0E2A498DAA0D4B83095ADA2288DCF5FEEA5DE0ECC15295431EFC436E79859448A70D76D286785CA31500D305179039F16C138AED1C9D4C2CDF9E
3-PP6807264-POA.pdf	A91B1AF8E94363B6A716B73469E526CB2A31443B08F243A69347D9F7AB9A6A8FEBF3C6A3F4EAE88F205D76BAADC3E0C21A68E243C0D78E51382E7901FC2FDA99
4-PP6807264-Authorization for internet communication.pdf	3F519423E55AA48D06D3995EDA7B0E04084262D835A332BC47433D8F6AEA6D191DC1227833725BBCD8989E5BCCB06EC3ED488054373F51B9C2C96C4BCE9FB05A
5-PP6807264-PDX.pdf	6A78C11B2704434CA763E0628C4BA48718EF1AC2F4D9086523EF4C758A921358D27A13657EB7016F4D76327BF177A45A3BFA3B4A37E3B759644D4CDE5888ECC1
6-0-PP6807264-IDS.pdf	84D463FD484782C6AA533130F28D8468C4B07313C9664258C8BE0DB8ABB12405DDE0B9A520DA6879E3ACC5B2600C4D93A415E717DE48B811052D94EA68C1D929
8-PP6807264-Drawing.pdf	8FD32DFD714E941F642F7D14B647E2E40DC3BD05BE2B499D81A840A663A22FF5680402AE6F523207E0990B97CC62AFF0DF424A9A48CBC40AE0AFBDB337054BD3
6-4-MDVAD Multimodal Diffusion for Video Anomaly Detection.pdf	C265BD2F7D39381849B81407A4D2C4B6C87E0589F3580ED1D68E80145FA1CB9DE728E79107F517E14B4CE614022EA1ABD2E3833E4D91EA6AD29644302D84C67D
6-1-KR10-2022-0088653-TR.pdf	9C206A4D9B42E05CAC9D7C132D493E64539EBA15C8FFCE8A5A1FC18CF34ABCAAF00A234B84F6B635C13AB563DD8A9A85E2DE637264E886E7322CD6A4E50C4171
9-PP6807264_Declaration_Assignment__.pdf	E81FC46BB4AFEB3C7D4CA58BD503736CF8B943977819BB8D770D4174EC26526583AD45590B4002BB4D61FF3D9D80BDA5BD9A417064EB61BCD911CBE3EF0BAFD2
6-2-JP2018-077479-TR____part_1.pdf	3146F74654E58A4B77EEF048105FD60881A9EC0CA5B73B5B1E2FA4D47CBCEDC6235464ACCB0CC568015B88D83DEA3ECEEFF0E09571B007738226A91DC082989E
6-2-JP2018-077479-TR____part_2.pdf	22374FA8665081B8A95347590579F043D28C30D3F789BDCC63315E2EA152696E2FCB925271B88697D1876DBF380C82B2FA222BC725C8AF8B54D65D6FED97ADA3
7-PP6807264 Specifications.pdf	9D79B02ED0921ECC154D3BF5686E03CB58A707B3CDDCBC45A27C487233BE9AE0D9742A3F0491F917934F968D6A700E935B25817CE4A0320DB7CA8FA881B9D9E4

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d)

and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

ELECTRONIC PAYMENT RECEIPT

APPLICATION #	RECEIPT DATE / TIME	ATTORNEY DOCKET #
19/078,589	**03/13/2025 12:33:59 PM Z ET**	**PP6807264**

Title of Invention

METHOD OF DETECTING VIDEO ANOMALY ON BASIS OF MULTIMODAL DIFFUSION AND DEVICE THEREFOR

Application Information

APPLICATION TYPE	Utility - Nonprovisional Application under 35 USC 111(a)	PATENT #	-
CONFIRMATION #	4817	FILED BY	BYUNGWOONG PARK
PATENT CENTER #	69565003	AUTHORIZED BY	-
CUSTOMER #	183715	FILING DATE	-
CORRESPONDENCE ADDRESS	-	FIRST NAMED INVENTOR	Sanghyun PARK

Payment Information

PAYMENT METHOD	PAYMENT TRANSACTION ID	PAYMENT AUTHORIZED BY
CARD / 7570	**E20253CC34348743**	**BYUNGWOONG PARK**

FEE CODE	DESCRIPTION	ITEM PRICE($)	QUANTITY	ITEM TOTAL($)
4011	BASIC FILING FEE- UTILITY	70.00	1	70.00
2111	UTILITY PATENT APPL. SEARCH FEE	308.00	1	308.00
2311	EXAMINATION OF ORIGINAL PATENT APPLICATION	352.00	1	352.00
			TOTAL AMOUNT:	**$730.00**

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d)

and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT ASSIGNMENT COVER SHEET

Assignment ID: 884810

Electronic Version v1.1
Stylesheet Version v1.2

SUBMISSION TYPE:	NEW ASSIGNMENT
NATURE OF CONVEYANCE:	ASSIGNMENT

CONVEYING PARTY DATA

Name	**Execution Date**
Sanghyun PARK	03/12/2025
Kijung LEE	03/12/2025
Youngwan JO	03/12/2025
Sunghyun AHN	03/12/2025

RECEIVING PARTY DATA

Name:	UIF (University Ind ustry Foundation), Yonsei University
Street Address:	50 Yonsei-ro, Seodaemun-gu
City:	Seoul
State/Country:	KOREA, REPUBLIC OF
Postal Code:	03722

PROPERTY NUMBERS Total:1

Property Type	**Number**
Application Number:	19078589

CORRESPONDENCE DATA

Fax Number: 7039954571
Phone: 3012460134
Email: bwp@zionip.com,pbwoong@hotmail.com
Correspondence will be sent to the e-mail address first; if that is unsuccessful, it will be sent using a fax number, if provided; if that is unsuccessful, it will be sent via US Mail.
Correspondent Name: Byungwoong Park
Address Line 1: 1717 K St NW, #879
Address Line 4: Washington , DISTRICT OF COLUMBIA 20006

ATTORNEY DOCKET NUMBER:	PP6807264
NAME OF SUBMITTER:	BYUNGWOONG PARK

Signature:	/BYUNGWOONG PARK/
Date:	03/13/2025
	This document serves as an Oath/Declaration (37 CFR 1.63).
TOTAL ATTACHMENTS: 1 source= PP6807264_Declaration_Assignment__.pdf	